

SEP

SECRETARÍA DE  
EDUCACIÓN PÚBLICA



TECNOLÓGICO NACIONAL DE MÉXICO  
Instituto Tecnológico de La Paz



# Investigación y Desarrollo en Robótica y Computación

**CIRC**  
2016

Coordinadores de la Edición  
Iliana Castro Liera  
Mario Cortés Larrinaga

ISBN: 978-607-97128-2-2



Esta página fue dejada en blanco intencionalmente



TECNOLÓGICO NACIONAL DE MÉXICO  
Instituto Tecnológico de La Paz



## Investigación y desarrollo en robótica y computación

D. R. © INSTITUTO TECNOLÓGICO DE LA PAZ  
Blvd. Forjadores de B.C.S. No. 4720  
Col. 8 de octubre  
C. P. 23080  
La Paz, Baja California Sur, México

Primera edición abril de 2016

ISBN: 978-607-97128-2-2



Reservados todos los derechos. Ninguna parte de este libro puede ser reproducida, archivada o transmitida, en cualquier sistema –electrónico, mecánico, de fotorreproducción, de almacenamiento en memoria o cualquier otro– sin hacerse acreedor a las sanciones establecidas en las leyes, salvo con el permiso escrito del titular de los derechos de la obra.

Diseño de la portada: Iliana Castro Liera

Esta página fue dejada en blanco intencionalmente

## COMITÉS

### **Coordinador General**

M.A.T.I. Luis Armando Cárdenas Florido - Instituto Tecnológico de La Paz

### **Coordinador del Programa Técnico**

Dr. Marco Antonio Castro Liera - Instituto Tecnológico de La Paz

### **Comité Revisor**

Dr. Francisco Herrera Fernández - UCLV Cuba

Dr. Alejandro Israel Barranco Gutiérrez - Instituto Tecnológico de Celaya

Dr. Víctor Adrián Santibáñez Dávila - Instituto Tecnológico de La Laguna

Dr. Ricardo Emmanuel Campa Cocom - Instituto Tecnológico de La Laguna

M.C. Edmundo Javier Ollervides Vázquez - Instituto Tecnológico de La Laguna

Dr. Eduardo Javier Moreno Valenzuela - CITEDI Tijuana

Dr. Eusebio Bugarín Carlos - Instituto Tecnológico de Ensenada

Dra. Ana Yaveni Aguilar Bustos - Instituto Tecnológico de Ensenada

Dr. Carlos Alberto Brizuela Rodríguez - CICESE

Dr. Guillermo Martínez Flores, CICIMAR - IPN

Dr. Jesús Alberto Sandoval Galarza - Instituto Tecnológico de La Paz

M.S.C. Iliana Castro Liera - Instituto Tecnológico de La Paz

Dr. Saúl Martínez Díaz - Instituto Tecnológico de La Paz

Dr. Israel Marcos Santillán Méndez - Instituto Tecnológico de La Paz

Dr. Marco Antonio Castro Liera - Instituto Tecnológico de La Paz

M.C. Jesús Antonio Castro - Instituto Tecnológico de La Paz

M.A.T.I. Luis Armando Cárdenas Florido - Instituto Tecnológico de La Paz

M.C. Juan Eduardo Machado Martínez - Instituto Tecnológico de La Paz

M.S.C. José Luís Gómez Torres - Instituto Tecnológico de La Paz

### **Coordinadores de la edición**

M.S.C. Iliana Castro Liera - Instituto Tecnológico de La Paz

Dr. Mario Cortés Larrinaga - Instituto Tecnológico de La Paz

**ÍNDICE**

Portada	
Comités	iv
Índice	v
Verificación Experimental de Osciladores Caóticos Booleanos	1
Diseño y construcción de un sistema de adquisición de datos para el monitoreo de corriente de un detector GEM	5
Diseño y construcción de un sistema de adquisición para el experimento de caminatas ópticas cuánticas con aplicación en cómputo cuántico	14
Sistema de sensorización no intrusiva en Entornos con Inteligencia Ambiental, a partir de casos de emergencia	23
Prototipo de sistema de telemetría de un satélite CanSat	29
Sincronización del bursting de la célula beta mediante redes complejas	34
Algoritmo de rastreo aplicado al tratamiento de niños con autismo	39
Reconocimiento de armas de fuego tipo pistola utilizando visión artificial	44
Contador de colonias bacterianas en placas Petri utilizando procesamiento digital de imágenes	49
Sistema mecatrónico para la identificación de dispositivos en tarjetas PCB usando visión por computadora	54
Implementación de un filtro de correlación no lineal en una GPU	59
Algoritmo paralelo sobre un sistema manycore para eliminar ruido en imágenes	63
Un algoritmo memético paralelo para TSP	70
Análisis de desempeño en la implementación de la DCT bidimensional en cómputo paralelo	75
Algoritmo de agrupamiento aplicado en la inteligencia de negocios usando GPGPU	81
Integración de CUDA C , OpenCV y R, para el procesamiento paralelo de imágenes de satélite	85
Diseño de un robot cartesiano para ser empleado en un sistema cartográfico de resolución controlada	89
Diseño e instrumentación de un sistema de recepción OFDMA en plataforma FPGA	98
Implementación de una Red Neuronal Artificial para la detección de personas en un Hardware Reconfigurable	104
Implementación del algoritmo de ICA en un DSP para la atenuación de ruido en llamadas telefónicas	108

Caracterización óptica de fibras ópticas sublongitud de onda	113
Nodo sensor infrarrojo para un arquitectura multicapa	117
Caracterización de micro-nano fibras ópticas fabricadas mediante un sistema mecatrónico desarrollado a partir de la tecnica flame brushing	123
Modelo de datos para el análisis de información georreferenciada sobre biodiversidad aplicado en un entorno de aplicaciones móviles	127
Análisis del Patrón de Comportamiento en la Participación Ciudadana de Elecciones Estatales	131
Implementación de algoritmos de control en un péndulo de transmisión directa	138
Diseño de un controlador LQR para un giroscopio de 3 grados de libertad	143
Implementación de algoritmos de control en tiempo real en la plataforma robótica Segway RMP-100	149
Control en orientacion de un Micro vehículo submarino mediante un sistema embebido	155
Diseño de control para evasión de obstáculos en un robot móvil omnidireccional utilizando Programación Genética	162
Path follower Algorithm for a Nao Humanoid Robot	168
Sobre el Cálculo del ZMP en la Estrategia Tobillo para el Balance de un Robot Humanoide	176
Implementación de un algoritmo de lógica difusa en el Software Matlab para un pre-diagnóstico de fallas en vehículos Volkswagen	182
Propuesta de Metodología para la Construcción de Buscadores Web Semánticos de un Contexto Especifico	188
Aplicación para el registro de resultados de estudios de Imagenología dentro del Expediente Clínico Electrónico mediante HL7	195
Sistema de Información para el Control de Acceso a Datos Personales	201
Protección de Imágenes a Color en el Dominio de la DWT Mediante Marca de Agua Invisible	208
Inserción de una Marca de Agua Visible Adaptativa basada en la Luminancia y Textura de la Imagen Digital Huésped a Color	215
Detección de la cobertura de manglar mediante el procesamiento de imágenes SPOT-5, Segmentación basada en NDVI y textura	220
Implementación de un algoritmo para la detección de polvo en imágenes MODIS	224
Construcción de modelo acústico en español de México para el área pediátrica y su implementación con CMU-SPHINX	228
Diseño e instrumentación de un sistema de monitoreo de pulso y temperatura en bebés vía Wi-Fi	233

# Verificación Experimental de Osciladores Caóticos Booleanos

Tonatiuh García\*, Jesús Muñoz\*, Luz del Carmen Gómez\*, and Víctor Rodolfo Gonzalez\*

\*Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, Puebla, México.

Email: tonaspiuck@gmail.com, jesus.munoz@correo.buap.mx, luz.gomez@correo.buap.mx y vicrodolfo.gonzalez@correo.buap.mx

**Resumen**—La computación no es el único uso para circuitos digitales. Otra aplicación de tales circuitos es la descripción de sistemas complejos por medio de redes booleanas. Este tipo de redes son vistas como una aproximación para entender grandes sistemas que son muy extensos como para ser modelados tomando en cuenta cada pequeño detalle. Generalmente existen dos aproximaciones para redes booleanas, aquellas que son síncronas y las que son autónomas. En este trabajo se enfoca en las redes booleanas del segundo tipo debido al complejo comportamiento que pueden presentar, específicamente caótico, y en los aspectos que hacen posible observar dicho comportamiento. Se muestra el resultado de la implementación electrónica de tres circuitos de redes booleanas autónomas que poseen un comportamiento caótico en tiempo continuo.

## I. INTRODUCCIÓN

Las redes booleanas fueron propuestas primeramente en 1965 por Walker y Ashby por el interés en sistemas complejos y popularizadas por Kauffman como un modelo para circuitos genéticos. Estas llamadas redes Kauffman fueron ampliadas por Ghil y Mullhaupt en 1984 con ecuaciones booleanas de retardo ya que tomaban en cuenta el tiempo que demora la propagación de las señales [1]. La descripción a partir de redes booleanas es comúnmente utilizada para modelar sistemas complejos, este tipo de sistemas generalmente exhiben un comportamiento con niveles de umbral, tienen múltiples retroalimentaciones en sus conexiones y las señales se propagan con diversos retardos en el tiempo. Ejemplos de su aplicación se pueden encontrar en biología, donde redes booleanas son usadas para modelar la interacción de los componentes en redes de regulación genética y también en redes neuronales. En geociencias las redes booleanas han sido utilizadas como una idealización de modelo climático para un rango amplio de escalas de tiempo, también para el modelado y predicción de terremotos, [2]-[4].

Generalmente se distinguen dos tipos de redes booleanas diferenciadas según el método de actualización de su estado: redes booleanas síncronas y redes booleanas autónomas. Las redes síncronas evolucionan en pasos de tiempo discretos, matemáticamente descritas por mapas iterados y experimentalmente realizadas con circuitos lógicos acoplados a un reloj que controla su actualización de estado, estas son del tipo Kauffman. Por otro lado, una red booleana autónoma (ABN, Autonomous Boolean Network) evoluciona en tiempo continuo, matemáticamente descrita por ecuaciones diferenciales o ecuaciones de retardo booleanas (BΔE, Boolean Delay Equations) y experimentalmente realizada con circuitos lógicos sin fuentes de reloj. El parámetro que controla la actualización de estado en una red booleana autónoma son los retardos

que se originan debido al tiempo de procesamiento de los componentes además del tiempo que tarda la transmisión de la señal para llegar de un punto a otro [3].

En este último tipo de redes booleanas se ha comprobado que bajo ciertas circunstancias presentan un comportamiento caótico. La teoría del caos demuestra que una relación determinística simple pero no lineal, puede generar una trayectoria dinámica extremadamente compleja que parece ser completamente aleatoria. Los tres indicadores del caos son: Un comportamiento de un sistema determinista, esto es, la evolución actual es consecuencia del estado previo del sistema, gran sensibilidad a condiciones iniciales y un comportamiento aperiódico a largo plazo.

En este trabajo se muestra el resultado de la implementación experimental de tres redes autónomas booleanas y su representación por medio de ecuaciones de retardo booleanas. La red 1 [5] contiene una compuerta XOR junto con un oscilador basado en una compuerta NOT y dos retardos en la propagación de las señales. La red 2 [1] está compuesta de tres compuertas, dos son del tipo XOR y una es XNOR, esta red contiene 6 retardos en la propagación de las señales y, por último, la red 3 [2] está basada en una compuerta XNOR de tres entradas, donde se retroalimenta a sí misma con tres retardos diferentes. En todos los casos son implementadas con compuertas lógicas comerciales de la familia 74HCXXX.

El documento está dividido de la siguiente manera; en sección II trata sobre las ecuaciones de retardo booleanas y su uso para el modelado de redes booleanas autónomas. En la siguiente sección se describen parámetros no lineales que conllevan a un comportamiento caótico en este tipo de sistemas, en la sección IV se presentan ejemplos de circuitos que muestran este tipo de comportamiento junto con el resultado de su implementación. Para finalizar en la sección V se dan las conclusiones del trabajo.

## II. REDES BOOLEANAS AUTÓNOMAS

Una red booleana autónoma, es un conjunto de nodos con valores binarios conectados por enlaces que tienen asociados retardos. Se establece un vector de variables booleanas  $\mathbf{x} = \{x_1, x_2, \dots, x_n\}$  que otorga una descripción cualitativa del sistema original. La dependencia del tiempo puede ser introducida mediante la prescripción de un conjunto de retardos  $t_{ij}$ , donde  $i = 1, \dots, n$ ,  $j = 1, \dots, n$ ,  $t_{ij} > 0$ , además  $t_{ij}$  es el tiempo que toma a  $x_j$  tener efecto en  $x_i$ . Por lo tanto para cada par de variables de estado existe un tiempo de retardo asociado. No necesariamente se cumple que  $t_{ij} = t_{ji}$ .



La relación que existe entre las variables booleanas es descrita mediante  $f_i : \mathbf{B}^n \rightarrow \mathbf{B}, i = 1, \dots, n$ , donde  $\mathbf{B} = \{0, 1\}$ , como se muestra:

$$\begin{aligned} x_1(t) &= f_1(x_1(t-t_{11}), x_2(t-t_{12}), \dots, x_n(t-t_{1n})), \\ x_2(t) &= f_2(x_1(t-t_{21}), x_2(t-t_{22}), \dots, x_n(t-t_{2n})), \\ &\vdots \\ x_n(t) &= f_n(x_1(t-t_{n1}), x_2(t-t_{n2}), \dots, x_n(t-t_{nn})), \end{aligned} \quad (1)$$

El sistema de ecuaciones de diferencias booleano, o  $B\Delta E$ , es más fácil de describir en términos de variables de memorización  $x_{ij}$ :

$$x_{ij}(t) = x_j(t-t_{ij}), \quad i = 1, \dots, n, \quad j = 1, \dots, n. \quad (2)$$

De esta manera, en términos de variables de memorización (1) se vuelve:

$$x_i(t) = f_i(x_{i1}(t), x_{i2}(t), \dots, x_{in}(t)) \quad i = 1, \dots, n. \quad (3)$$

La ecuación (3) establece que en el tiempo  $t$  el estado  $x_i$  es determinado por las variables de memorización  $x_{i1}, x_{i2}, \dots, x_{in}$ , o equivalentemente, por los estados  $x_j$  en tiempos  $t-t_{ij}, j = 1, \dots, n$ .

Con el fin de resolver el sistema de ecuaciones (1) necesitamos definir el valor de las variables booleanas en el intervalo:

$$x_i(t) = x_i^0 = (t) \quad \text{para } t_0 - \tau \leq t < t_0, \quad i = 1, \dots, n, \quad (4)$$

donde  $\tau = \max\{t_{ij}\}$ . Este  $\tau$  es llamado la dimensión de memoria del sistema debido a que el estado del sistema en el tiempo  $t$  depende solamente de los estados  $\{x_i(t-t')\}, 0 < t' \leq \tau$ . Después de estos preliminares se puede establecer el siguiente teorema de existencia y unicidad de las soluciones [3].

*Teorema 1.-* Siendo  $\{x_i^0(t)\}$  los datos iniciales con saltos de 0 a 1 o viceversa en un número finito de puntos sobre el intervalo  $t_0 - \tau \leq t < t_0$ . Entonces el sistema (1) tiene solución única para todo  $t \geq t_0$ , para retardos arbitrarios  $t_{ij} > 0$ .

### III. CAOS BOOLEANO

De la misma manera que no cualquier sistema dinámico continuo puede exhibir caos, no cualquier tipo de red booleana puede mostrar un comportamiento caótico. A continuación se exponen los parámetros que se involucran para la obtención del fenómeno caótico en una red autónoma booleana [1], [3], [6].

- La asimetría entre estados lógicos. Básicamente es tomar en cuenta la no idealidad en la diferencia del tiempo de subida y tiempo de bajada en un cambio de estado booleano.

En [1] y [3] se muestra un circuito ideal compuesto de una compuerta XOR y dos lazos de retroalimentación con retardos

inconmensurables, lo que provoca que el sistema evolucione a estados dinámicos con anchos de pulso lo más pequeños posibles. Este comportamiento complejo conlleva a lo que llaman una catástrofe ultravioleta donde los pulsos se vuelven cada vez más densos en el tiempo y la anchura de los pulsos tiende a cero. Cabe aclarar que este fenómeno nunca ocurriría a nivel experimental debido a las limitaciones de procesamiento en las compuertas lógicas, de la misma manera tiene sus restricciones en simulación debido a que, según el valor de los retardos, los pulsos pueden crearse cada vez más juntos entre sí, incluso más que la exactitud del algoritmo [3], [6].

Lo anterior nos lleva al segundo efecto no lineal:

- Este fenómeno es llamado rechazo a pulso corto (short-pulse rejection). Dicho fenómeno establece que los nodos en la red no pueden procesar pulsos de una pequeña duración arbitraria; pulsos más pequeños que la duración de corte  $\tau_{spr}$  son filtrados de tal manera que no tienen efecto en el nodo.

El último efecto no lineal involucrado para la observación del fenómeno caótico es quizá el más significativo. Este efecto involucra la historia de la señal cuando una nueva transición ocurre y el tiempo que ha pasado desde su anterior cambio de nivel [7].

- El efecto de degradación. Esta no linealidad toma lugar cuando surge una transición en la entrada y la compuerta aún no ha cambiado completamente de nivel después que la transición anterior fuera completamente propagada.

Existen otros factores que se deben tomar en cuenta para poder obtener un comportamiento caótico dentro de una red booleana, por ejemplo en [6] mencionan que la red no debe tener puntos fijos, es decir, estados en donde el sistema permanezca indefinidamente hasta que se aplique una perturbación. Otro parámetro es el tipo de operación lógica que ejercen los nodos, principalmente se ha mostrado el uso y combinación de compuertas del tipo XOR y XNOR para poder cumplir el objetivo. Sin embargo, el factor clave entre estos sistemas es el valor de los retardos en la propagación de la señal, estos retardos deben ser inconmensurables para tener una mayor probabilidad de éxito, se ha demostrado que un sistema que presenta el fenómeno caótico puede caer en un comportamiento completamente periódico al cambiar un poco el valor de algún retardo dentro del sistema [3].

### IV. OSCILADORES CAÓTICOS BOOLEANOS

En este trabajo se han recompilado tres circuitos booleanos que presentan un comportamiento caótico, en la Figura 1 se muestra cada sistema booleano que se ha implementado. Las ecuaciones (5-7) son las que rigen la dinámica de cada circuito respectivamente [1], [2], [5].

En la imagen  $\tau_{i,j}$  es el retardo involucrado en la propagación de la señal, es el tiempo que tarda la salida del punto  $j$  en tener efecto sobre la entrada  $i$ . En el caso del circuito c) se podría ver como una compuerta XNOR de tres entradas retroalimentadas donde cada lazo de retroalimentación tiene implícito un retardo, recalando que deben ser de valor inconmensurable diferente.

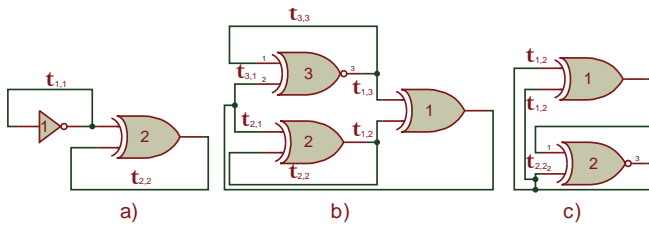


Figura 1. Osciladores Caoticos Booleanos implementados en este trabajo

Ecuaciones para el circuito presentado en la Figura 1-a):

$$\begin{aligned} x_1(t) &= \neg x_1(t - \tau_{1,1}) \\ x_2(t) &= x_1(t) \oplus x_2(t - \tau_{2,2}) \end{aligned} \quad (5)$$

Ecuaciones para el circuito presentado en la Figura 1-b):

$$\begin{aligned} x_1(t) &= x_3(t - \tau_{1,3}) \oplus x_2(t - \tau_{1,2}) \\ x_2(t) &= x_2(t - \tau_{2,2}) \oplus x_1(t - \tau_{2,1}) \\ x_3(t) &= \neg(x_1(t - \tau_{3,3}) \oplus x_2(t - \tau_{3,1})) \end{aligned} \quad (6)$$

Ecuaciones para el circuito presentado en la Figura 1-c):

$$\begin{aligned} x_1(t) &= x_2(t - \tau_{1,2}) \oplus x_2(t - \tau_{1,2}) \\ x_2(t) &= \neg x_1(t) \oplus x_2(t - \tau_{2,2}) \end{aligned} \quad (7)$$

donde el símbolo  $\neg$  significa la operación lógica NOT y el símbolo  $\oplus$  es la operación lógica XOR.

Como se menciono anteriormente, el valor de los retardos es un parámetro crucial para poder obtener o no un comportamiento caótico. Estos parámetros fueron ajustados durante el proceso de implementación debido a que cada circuito fue construido con compuertas comerciales y no con la misma tecnología de donde se obtuvieron los retardos iniciales. Los resultados de la implementación de cada circuito se muestran en las Figuras 2-7.

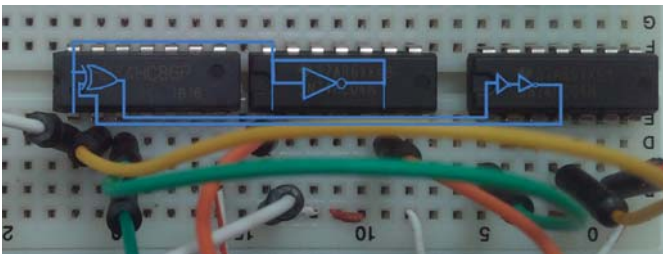


Figura 2. Implementación del circuito de la Figura 1-a.

Para verificar cuantitativamente el comportamiento caótico de los sistemas se calculo el exponente de Lyapunov siguiendo el algoritmo de [8]. El exponente de Lyapunov mide que tan sensible es la señal a las condiciones iniciales, es decir el grado de divergencia entre dos trayectorias cercanas conforme transcurre el tiempo. Un valor positivo del exponente de Lyapunov es prueba de la existencia de un comportamiento caótico. Los resultados se muestran en la Figura 8.

En la Figura 8 se muestran cuatro graficas de tres circuitos, esto es debido a que en el circuito 3 se encontraron dos

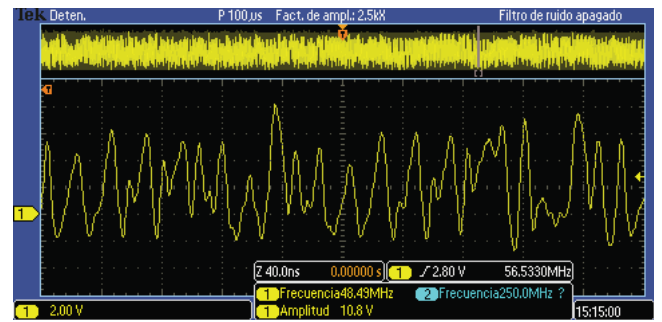


Figura 3. Resultado de la implementación del circuito de la Figura 1-a.

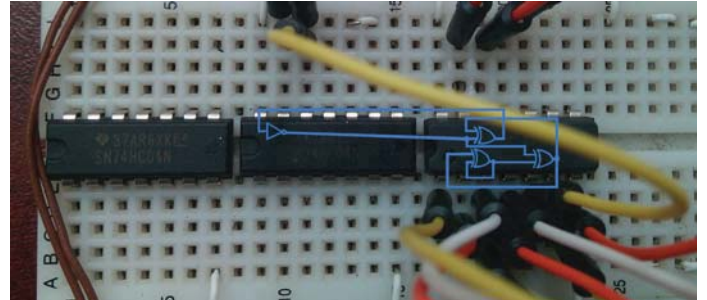


Figura 4. Implementación del circuito de la Figura 1-b.



Figura 5. Resultado de la implementación del circuito de la Figura 1-b.

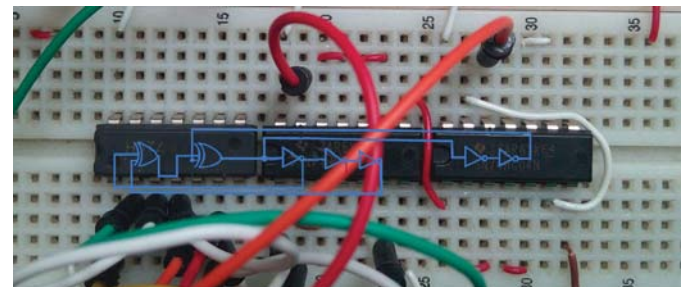


Figura 6. Implementación del circuito de la Figura 1-c.

configuraciones en las que presenta caos dependiendo del valor de los retardos en la propagación de la señal. El valor del exponente de Lyapunov es 0.1299, 0.1794, 0.1396 y 0.1162, para el circuito 1, circuito 2 y circuito 3 a-b, respectivamente.

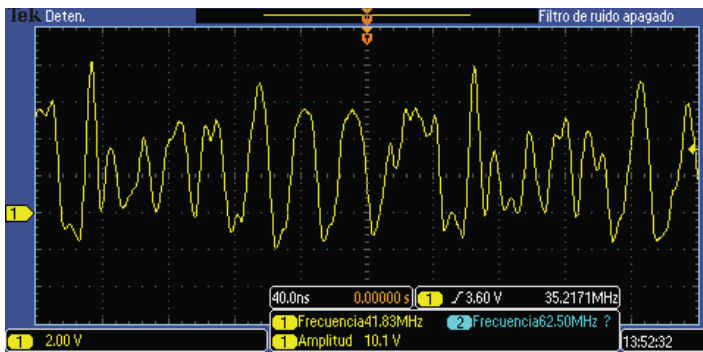


Figura 7. Resultado de la implementación del circuito de la Figura 1-c.

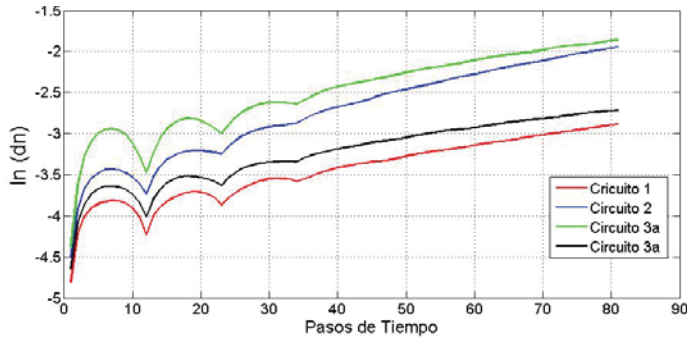


Figura 8. Logaritmo Natural de la divergencia para cada sistema.

## V. CONCLUSION

Se han presentado tres osciladores caóticos booleanos los cuales se describen por medio de ecuaciones de retardo. Se ha demostrado experimentalmente que en los tres casos es posible obtener comportamiento caótico. Se ha calculado el exponente máximo de Lyapunov para asegurar el fenómeno caótico. Los circuitos resultantes han sido implementados con compuertas XOR, XNOR y NOT de uso comercial.

## REFERENCIAS

- [1] Hugo L. D. de S. Cavalcante, Daniel J. Gauthier, Joshua E. S. Socolar, Rui Zhang, On the Origin of Chaos in Autonomous Boolean Networks, Duke University, Department of Physics and Center for Nonlinear and Complex Systems, Durham, NC 27708 USA
- [2] David P. Rosin, Dynamics of Complex Autonomous Boolean Networks, Springer Theses, 2015.
- [3] Dee, D. & Ghil, M. 1984 Boolean Difference Equations, I: Formulation and Dynamic Behavior. SIAM J. Appl. Math. 44, 111-126.
- [4] Ghil, M., & Mullhaupt, A.P. 1985 Boolean delay equations II: Periodic and aperiodic solutions. J. Stat. Phys. 41, 125-173.
- [5] Myunghwan Park, John C. Rodgers and Daniel P. Lathrop, True Random Number Generation based on CMOS Boolean Chaotic Oscillator, IEEE Transactions on Circuits and Systems, Vol. 1, No. 1. July 2013.
- [6] Zhang, R., Cavalcante, H.L.D.S., Gao, Z., Gautier, D.J., Socolar, J.E.S, Adams, M.M. & Lathrop, D.P. 2009 Boolean Chaos, to appear as a rapid communication in Physical Review E.
- [7] M.J. Bellido Díaz, J. Juan Chico, A. J. Acosta, M. Valencia and J. L. Huertas, Logical modelling of delay degradation effect in static CMOS gates, IEE Proc-Circuits Devices Syst., Vol 147, No. 2, April 2000.
- [8] M. Rosenstein, J. Collins, and C. De Luca, A practical method for calculating largest Lyapunov exponents from small data sets, Physica D: Nonlinear Phenomena, vol. 65, no. 1-2, pp.117-134,1993.

# ***DISEÑO Y CONSTRUCCION DE UN SISTEMA DE ADQUISICION DE DATOS PARA EL MONITOREO DE CORRIENTE DE UN DETECTOR GEM***

Ricardo Alberto De Gante Bautista.  
Benemérita Universidad Autónoma de Puebla.  
Facultad de Ciencias de la Electrónica, México.  
Email: ricardoalberto.degante@alumno.buap.mx

Sergio Vergara Limon.  
Benemérita Universidad Autónoma de Puebla  
Facultad de Ciencias de la Electrónica, México.  
Email: svergara2@hotmail.com

Guy Paic  
Universidad Nacional Autónoma de México  
Instituto de Ciencias Nucleares, México.  
Email: Guy.Paic@cern.ch

M. Aurora Diozcora Vargas Treviño.  
Benemérita Universidad Autónoma de Puebla.  
Facultad de Ciencias de la Electrónica, México.  
Email: auroravargast@hotmail.com

Fernando Reyes Cortes.  
Benemérita Universidad Autónoma de Puebla.  
Facultad de Ciencias de la Electrónica, México.  
Email: fernando.reyes@correo.buap.mx

**Resumen:** ALICE es el gran experimento colisionador de iones parte del programa del LHC (Large Hadron Collider) del CERN situado en la frontera de Suiza y Francia. Actualmente el espectrómetro de ALICE está instalado en el punto 2 que se encuentra en el lado Francés del CERN. El espectrómetro de ALICE consiste en 16 detectores, uno de ellos es el detector cámara de proyección de tiempo (TPC, Time Projection Chamber). El principal objetivo de este detector es determinar el momento y tipo de partícula entregando una imagen tridimensional de la partícula detectada. Se han desarrollado nuevos detectores para ser instalados en la TPC, denominados GEM (GAS ELECTRON MULTIPLIER), los cuales permiten detectar partículas o electrones, los GEM amplifican la partícula detectada para poder procesar después la información, como característica importante los GEM trabajan con alto voltaje y en función al proceso de amplificación de las partículas o electrones se producen descargas eléctricas aleatoriamente durante la operación de la TPC. En este trabajo se describe una idea general del funcionamiento de los detectores GEM, se ha diseñado e instrumentado un sistema de monitoreo de corriente que funcione con el alto voltaje, que permita visualizar y medir las descargas en este detector durante su operación, sin interferir con su funcionamiento normal. Con la información que se obtenga se considera determinar el desgaste y tiempo de vida útil del detector, en función a la información obtenida, esta

información se mostrara y almacenara de forma inalámbrica en una PC.

## **Abstract:**

ALICE is the great collider ions experiment, is part of the LHC (Large Hadron Collider) CERN program in located between Switzerland and France. At present the ALICE spectrometer was installed in the point 2 which is located at the CERN French side. The ALICE spectrometer consist in 16 detectors, one of them is the TPC (Time Projection Chamber). The main objective of this detector is determine the moment and type of particle, providing a three-dimensional image of the particle detected. Developing new detectors to be installed in the TPC, the name of these detectors is GEM (Gas Electron Multiplier), detect particles or electrons, GEM amplifies the detected particle, for later process the information, as important characteristic, the GEM works with high voltage and according to the process of amplification of the particles or electrons, electric shock occurring during operation of the TPC. In this paper describe a general idea of the GEM detector function, was designed and instrument a monitoring system current, this system work with high voltage and permit measure electrical shocks in the detector during normal operation. Whit the information obtained is determined wear and lifetime of the detector. The information display and stored wirelessly to a PC.

## I. Introducción.

ALICE (A Large Ion Collider Experiment) es el gran experimento colisionador de iones parte del programa del LHC (Large Hadron Collider) del CERN en Suiza. En la actualidad el espectrómetro de ALICE se está instalando en el Punto 2 que se encuentra en el lado Francés del CERN [2], ver Fig. 1.

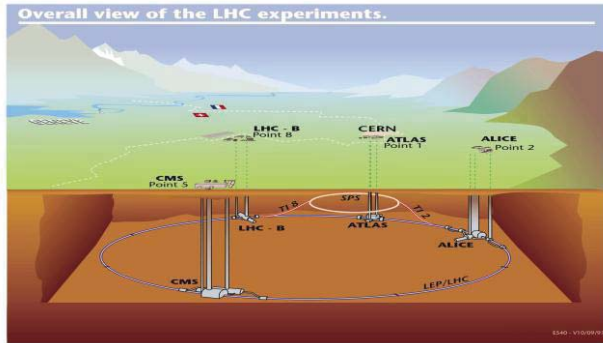


Fig. 1: Programa del LHC del CERN.

El detector ALICE ha sido especialmente concebido para analizar un estado particular de la materia, denominado plasma de quarks y de gluones, que se cree existió justo después de la creación del Universo. El espectrómetro de ALICE consiste en 16 detectores, uno de ellos es la TPC (Time Projection Chamber) [2] Fig.2.

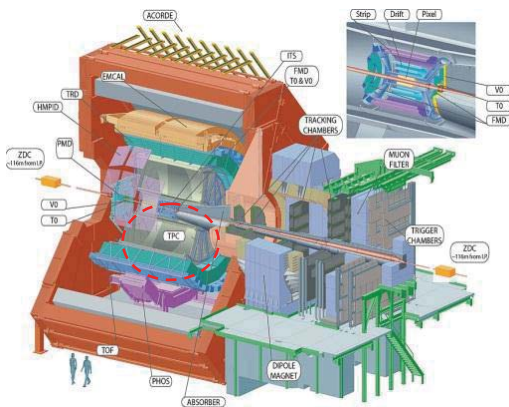


Fig. 2: Espectrómetro de ALICE.

## II. TPC (Time Projection Chamber) de ALICE.

TPC introdujo un nuevo nivel de rendimiento que les permite a los físicos a analizar las colisiones de partículas en tres dimensiones [3]. Cuando una partícula cargada, tal como un positrón, viaja a través de una sustancia como el aire, se ionizan átomos que cruzan su trayectoria, golpeándose liberarán algunos de sus electrones, no se puede observar la partícula ionizante directamente, pero se pueden utilizar detectores de partículas para amplificar y medir la estela de electrones libres. Esos electrones libres sirven

como sustituto de la trayectoria de la partícula ionizante, que permiten ver el tipo de partícula y su momento.

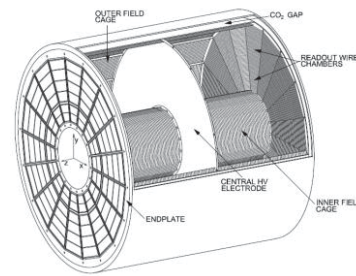


Fig. 3: TPC Proyecto ALICE.

La TPC ofrece una imagen 3D de la ionización depositada en un volumen de un gas o líquido. Actúa como una cámara de burbujas. La localización 3D de la TPC es extremadamente útil en el seguimiento de las partículas cargadas en un entorno de alta densidad y para identificar partículas a través de su pérdida de energía de ionización. Los detectores en la TPC son llamados GEM [4].

## GEM (GAS ELECTRON MULTIPLIER)

La actividad reciente de investigación se ha centrado en el desarrollo de la tecnología del GEM, inventado en 1997 por Fabio Sauli.

Un GEM como se muestra en la figura4 está hecho de una placa (50 micras) de lámina de Kapton, un revestimiento de cobre en cada lado, perforado superficialmente, cada perforación actúa como un canal de multiplicación de electrones. Cada agujero tiene una estructura bi-cónica con diámetro externo de 50 micras y un paso de 140 micras, como en la Fig. 5 se ilustra [4].

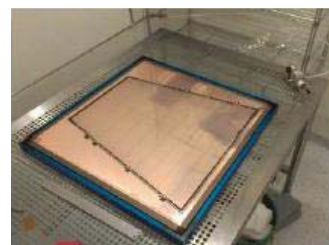


Fig. 4: GEM.

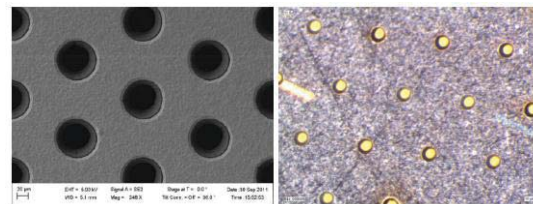


Fig. 5: Fotografía de un microscopio electrónico de un GEM.

El funcionamiento GEM actual cuando una partícula cargada o un fotón que interactúan en el gas de la región del drift producen una ionización racimos de iones positivos y electrones como se observa en la figura 6. Debido al campo eléctrico aplicado, el grupo de iones positivos del drift va hacia el cátodo, mientras que el grupo de electrones va hacia los agujeros de la primera lámina de GEM.

Aquí la carga se multiplica y mientras la carga amplificada hacia el segundo GEM se frustra, un gran porcentaje de los iones positivos producidos en los agujeros es capturado por el electrodo superior GEM.

Así canal multiplicación se libera rápidamente y grupos de electrones se multiplican. La ganancia del detector es proporcional la suma de los voltajes de GEM [1].

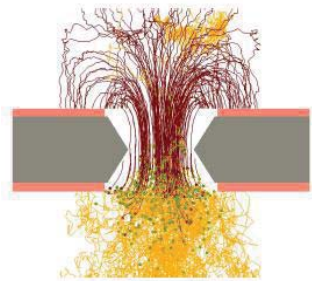


Fig. 6: Funcionamiento GEM.

**Prestaciones del detector GEM**

Un único detector GEM consiste en una región de conversión y “drift”, una lámina de GEM, una región de inducción y una placa de circuito impreso (PCB) para la recolección de electrones [1]. Esta configuración se muestra en la Fig. 7.

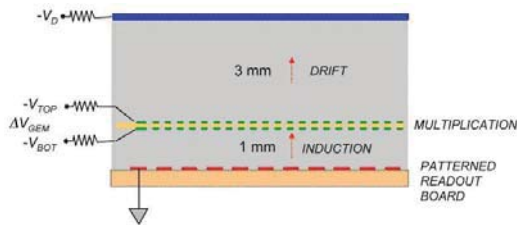


Fig. 7: Conexión de un GEM simple

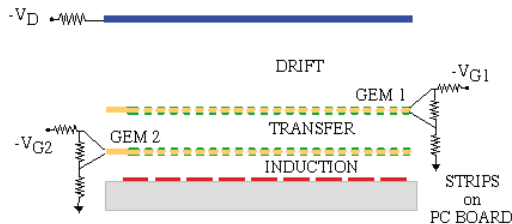


Fig. 8: Esquemático que muestra la conexión de un GEM doble.

La amplificación de avalancha proporcional se ha observado en una amplia gama de gases y presiones. La

ganancia efectiva de la estructura, se define como la relación recogida de carga primaria (una fracción de la carga es recogida por los electrodos GEM), puede llegar a diez mil en la mayoría de los gases. En las Fig. 9 se muestra la gráfica comparativa de respuesta del detector en relación al gas aplicado. Mostrando la ganancia obtenida (Effective Gain) en relación a el voltaje aplicado ( $\Delta V$ ) de un solo GEM [1].

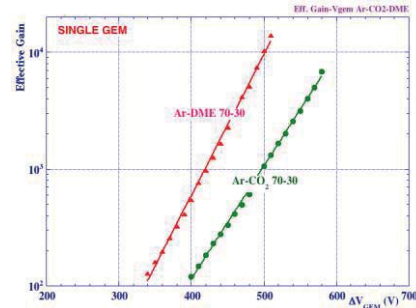


Fig.9: Relación de comparación entre la respuesta de amplificación de un GEM de una capa, respuesta con dos gases diferentes.

En las Fig. 10 se muestra la gráfica comparativa de respuesta del detector en relación al gas aplicado. Mostrando la ganancia obtenida (Effective Gain) en relación a el voltaje aplicado ( $\Delta V$ ) de un GEM doble. Se puede observar que la ganancia final en esta configuración es la suma de las ganancias de cada GEM [1].

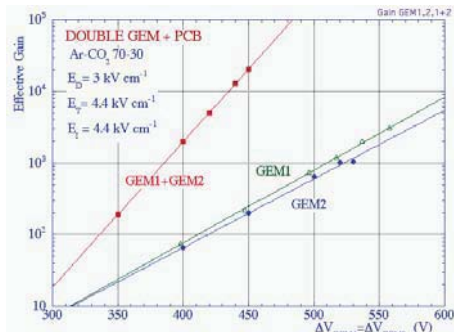


Fig. 10: Ganancias de un GEM doble en comparativa con dos GEM's separados.

Entre mayor sea el número de detectores GEM colocados en serie y debido a el voltaje aplicado en cada una de las capas de estos detectores la posibilidad de que exista una descarga es mayor. De la misma forma entre mas GEM sean colocados en conjunto con el nivel de voltaje aplicado mayor será la amplificación de los eventos.

**III. Sistema de monitoreo de corriente del detector GEM**

La electrónica fue diseñada y será instrumentada para conectarlo a los detectores GEM's con la finalidad de monitorear la corriente de consumo y poder obtener

información del estado de los detectores, determinar el tiempo que pueden estar en operación. Esta información será procesada en la tarjeta de desarrollo FPGA desarrollada por el grupo de robótica para posteriormente mostrar los resultados en una PC de manera inalámbrica transmitiéndola vía WiFi de acuerdo al diagrama de la Fig. 11.

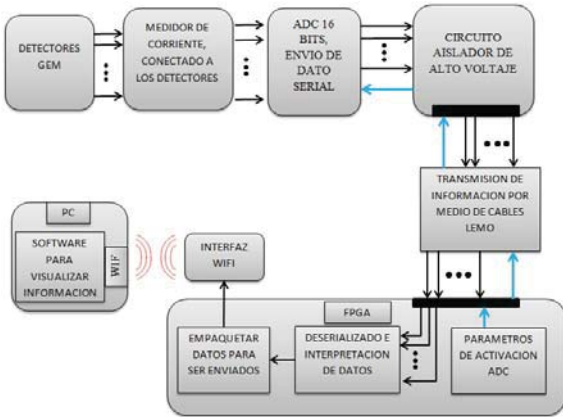


Fig. 11: Diagrama a bloques general del sistema propuesto.

#### Conexión del Sistema de monitoreo (Amperímetro)

El amperímetro se conecta al GEM en paralelo a la resistencia de 100 MΩ. Con un valor de resistencia alta con la finalidad que al tener una alta resistencia no afecte el funcionamiento del detector GEM.

#### ADC 16 Bits

Las magnitudes físicas son analógicas y normalmente el procesado de señal se realiza de forma digital, para ello la señal de corriente analógica debe ser convertida a un dato digital por medio de un convertidor analógico digital (ADC) de 16 bits. Esta resolución nos permitirá fidelidad en los datos muestreados. Para la transmisión de los datos digitales generados por el ADC se enviara esta información de manera serial.

#### Circuito aislador de alto voltaje

En este bloque como los medidores de corriente deben estar acoplados a la fuente de alto voltaje, se utilizara un circuito de aislamiento para evitar que el alto voltaje dañe los circuitos de bajo voltaje. Por lo anterior se utilizara un circuito optoelectrónico que tanga la capacidad de aislar hasta 10,000 Voltios. De esta forma se conectara la información entregada de forma serial a un circuito optoelectrónico que aislara los circuitos conectados al alto voltaje y los circuitos de bajo voltaje, protegiendo de esta forma la electrónica más sensible de posibles descargas eléctricas producidas por el alto voltaje.

#### FPGA

La información serial del canal será recibida por el FPGA, dentro del cual se instrumentaran los parámetros de configuración del ADC. También en el FPGA se instrumentara un deserializador para interpretar el dato recibido, para posterior mente empaquetar y enviar esta información vía WIFI una computadora, donde podrán ser almacenados y procesados.

#### PC

Con la información recibida por medio de la comunicación inalámbrica de la conexión WIFI, se desarrollara el software de interface visual para almacenar y desplegar la información.

#### IV. Hardware del sistema de monitoreo.

El diagrama de la Fig.12 muestra el hardware que se implementara para el desarrollo de un canal del sistema de monitoreo. Para la instrumentación de este diagrama se realizaron pruebas de funcionamiento de algunos componentes al alto voltaje, además que se caracterizó la respuesta de operación de los componentes.

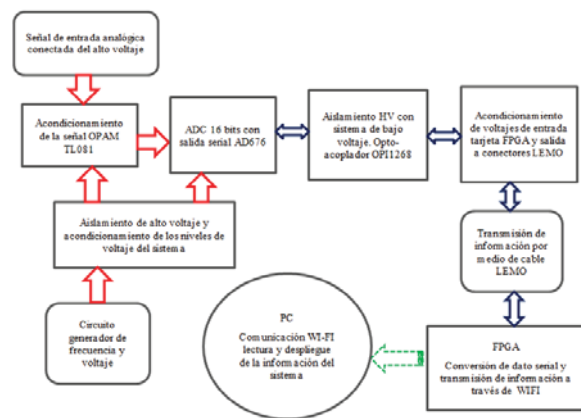


Fig. 12: Diagrama a bloques del primer sistema desarrollado.

#### Aislamiento de alto voltaje y acondicionamiento de los niveles de voltaje del sistema

Debido a la operación del sistema de monitoreo y su interacción con el alto voltaje, se implementa un transformador con un voltaje de aislamiento de 6Kv, con esto podemos aislar las posibles descargas de los detectores, con la implementación de este transformador, se regularan los voltajes que se requieren para la operación del sistema de acondicionamiento de señal y el ADC.

Dentro de las características del transformador seleccionado se tiene como ventaja el tamaño físico del componente.

**Caracterización transformador:**

En la Fig. 13 se muestra el transformador elegido para instrumentarlo en el sistema.



Fig.13: Transformador 76250ENC.

Se encontro el tipo de onda y frecuencia de operacion del transformador 76250ENC [5], se realizó un barrido de frecuencias para encontrar el óptimo funcionamiento del dispositivo como se muestra en la Fig. 14.

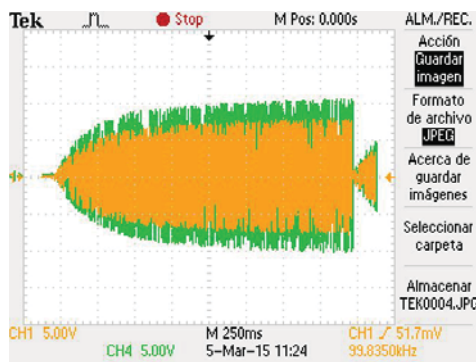


Fig. 14: Barrido de frecuencia del 76250ENC.

En relación con los resultados obtenidos, la frecuencia seleccionada fue en donde presenta un comportamiento estable, esta es de 33-35 KHZ con una señal de entrada cuadrada.

Una vez obtenida la frecuencia de operación del transformador se revisó el comportamiento y salida de los niveles de voltaje de los reguladores, obteniendo los niveles de voltaje requeridos (+5 y -5 Volts) para la operación de la electrónica en relación a las características de consumo y operación de los componentes seleccionados para la instrumentación del sistema de monitoreo.

**Circuito generador de frecuencia y voltaje.**

En la Fig. 15 se muestra la imagen de la tarjeta que se desarrolló para generar la señal requerida para el funcionamiento del transformador.

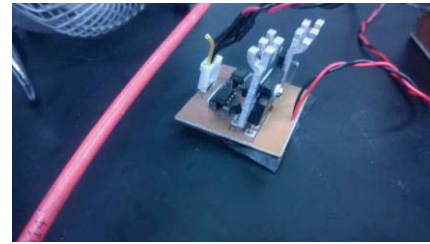


Fig. 15: Tarjeta generadora de señal.

Se implementa de la tarjeta generadora y se realizan pruebas de funcionamiento del transformador y niveles de voltaje. En la Fig. 16 se muestra la respuesta obtenida por el transformador con la señal generada con la tarjeta

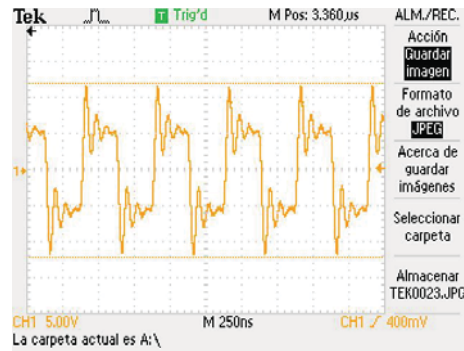


Fig. 16: Respuesta Transformador.

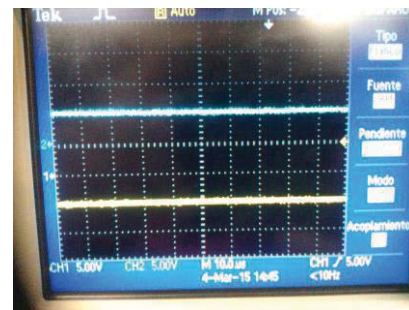


Fig. 17: Niveles de voltaje

En la Fig. 17 se muestran los niveles de voltajes obtenidos (+5 y -5 Volts) para la operación del sistema de monitoreo en relación a la respuesta de reguladores.

**Acondicionamiento de señal**

Para la conexión del sistema de monitoreo con el amperímetro se seleccionó el amplificador operacional TL081 [8] y se realizaron pruebas de funcionamiento, para esto se fabricó una tarjeta de prueba como se muestra en la Fig. 18.





Fig. 18: Tarjeta de prueba para OPAM TL081.

Se realizaron pruebas del OPAM TL081 con los niveles obtenidos con el transformador y los reguladores, se encontro que el amplificador tiene un comportamiento estable y un bajo nivel de offset, en comparativa con otros amplificadores operacionales.

### Convertidor analógico digital (ADC)

La conversión de la señal analogica se realizara con un ADC de 16 bits, AD676 [6] que tiene la capacidad de 100 kSPS, con un votaje de operación de 5 VDC, una salida serial de la conversión, habria que añadir que el reloj master del ADC y el inicio de conversión puede ser modificada por señales externas al dispositivo, de esta manera es posible configurar el ADC de manera remota.

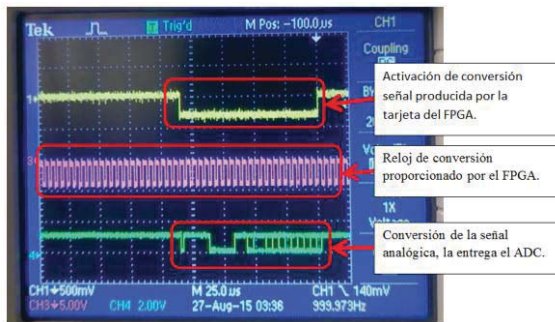


Fig. 19: Respuesta del ADC.

En la Fig. 19 se realizo la caracterizacion del ADC en funcion a la respuesta con las señales enviadas por el FPGA, y con el protocolo de activacion del componente, se observo que el comportamiento de la salida serial es deacuerdo a lo reportado por el fabricante.

### Aislamiento del alto voltaje

Para el aislamiento de la electronica del monitor de corriente que interactua con el alto voltaje de electronica que trabaja con bajo voltaja con bajo voltaje se seleccionaron los opto acompladores OPI1268 [7], ya que estos componentes tiene un una respuesta comprobada de alta velocidad de 2Mbit/s, un voltaje de

aislamiento de 16 kV y tiene una respuesta compatible TTL.

Acondicionamiento de voltajes para FPGA, se ocuáran buffer ya que se necesitan niveles de voltaje de 3.3VDC para enviar por el conector LEMO y de entrada al FPGA.

### Desarrollo y fabricacion de la PCB del sistema de monitoreo de corriente

Se darrolla en base a las puebas relizadas y caracterizacion de los componentes las PCB, para poder fabricarla, tomando en cuenta diversas estrategias de diseño para la disminucion de ruido en el la tarjeta. En la Fig.20 se muestra el diseño de la PCB.

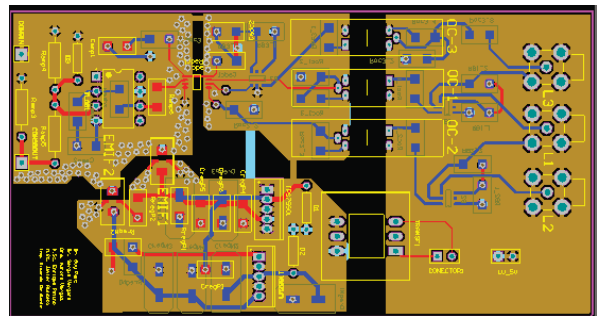


Fig. 20: PCB sistema de monitoreo.

Como se muestra en la Fig. 21 y Fig. 22 se fabrico la tarjeta y se intrumento para realizar las pruebas de funcionamiento.

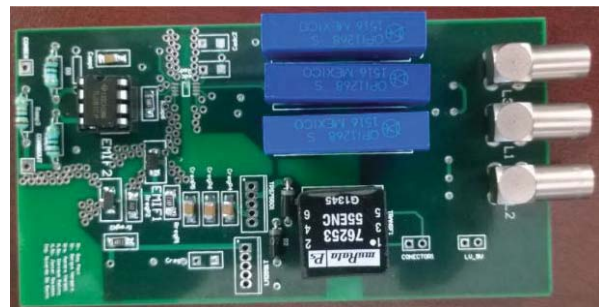


Fig. 21: Tarjeta fabricada PCB vista superior.

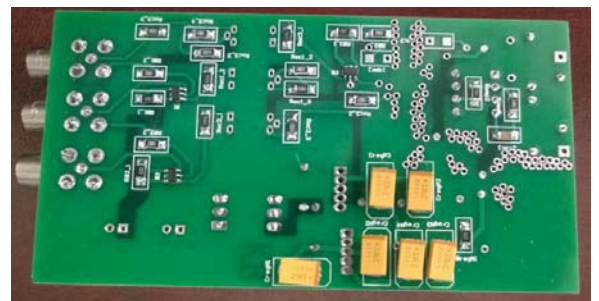


Fig. 22: Tarjeta fabricada PCB vista inferior.

**FPGA**

Se ocupara la tarjeta de FPGA desarrollada en el grupo de robotica de la Fig. 23 que cuenta con un dispositivo FPGA Cyclone III de 10,320 elementos logicos, 46 bloques de memoria, 2 PLLs y 182 pines de entrada/salida. En la tarjeta de desarrollo se realizara la deserializacion de los datos e interpretacion de estos valores para enpaquetarlos y enviarlos via WiFi, ademas se generaran las condiciones y señales para la activacion de ADC. Cabe mencionar que se ocupara el modulo de conexión WiFi para enviar informacion a la PC.



Fig. 23: Tarjeta con los conectores LEMO y WI-FI

**V. Firmware del sistema de monitoreo**

En la tarjeta de desarrollo del FPGA se desarrolla el Firmware para enviar el reloj de conversión de ADC y el bit de inicio de conversión, ademas la deserializacion del dato que envia el ADC. En la Fig. 24 se presenta el diagrama a bloques del firmware.

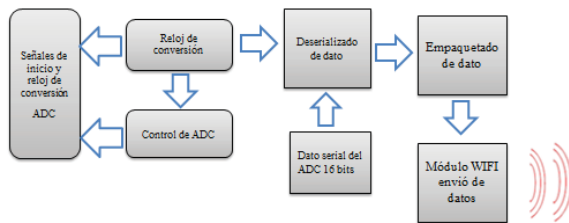


Fig. 24: Diagrama a bloques Firmware.

En la figura 25 se muestra el dato deserializado en el FPGA, se puede observar en círculo color azul el dato capturado o recibido por el analizador lógico del FPGA y por ende es el dato que lee la tarjeta de desarrollo del FPGA .

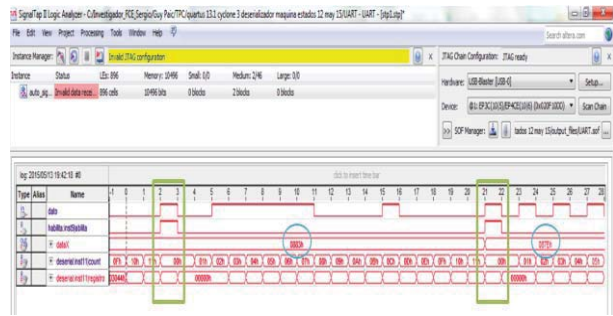


Fig. 25: Analizador lógico FPGA dato serial (SIGNAL TAP II LOGIC ANALYZER)

**VI. Software del sistema de monitoreo**

Como se muestra en la Fig. 26 para el desarrollo del software se tiene que realizar la conexión WIFI, recibir los paquetes de datos que se envían procedentes del FPGA, posteriormente se decodifican y verifica el orden de los datos recibidos para poder interpretar y desplegar la información en la pantalla, graficando los resultados obtenidos. En todo momento se verifica el estado de conexión inalámbrica WIFI y de ser necesario se reinicia y reconecta automáticamente.

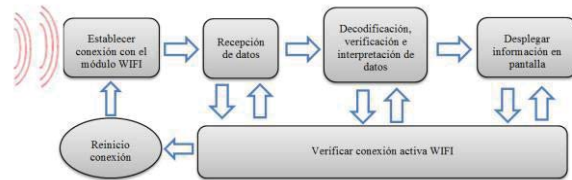


Fig. 26: Diagrama a bloques software.

Una vez instrumentada y desarrollado el software se realizó la prueba de comunicación WiFi del FPGA con la PC. Para realizar la comunicación con la PC es necesario configurar el enlace se ocupó el software UTF-8 Tera Term Pro. Con la comunicación realizada y configurado el sistema wi-fi se enlazara con el software para el despliegue de la información obtenida por la interface, con esto validaremos que la información y valores que ingresamos en la entrada de la interface están llegando a su destino.

Se desplegará la información obtenida por el amperímetro como se muestra en la Fig. 27 en el graficador del lado izquierdo en una escala de  $\mu A$ , del lado derecho se encuentra un graficador para verificar el voltaje en el ADC para verificar el comportamiento en una escala de mV. Para verificar el los niveles de voltaje obtenidos y el correcto funcionamiento, se colocaron indicadores para mostrar que la conexión wi-fi esta activa, verificar que el dato llegue en base a un byte adicional.

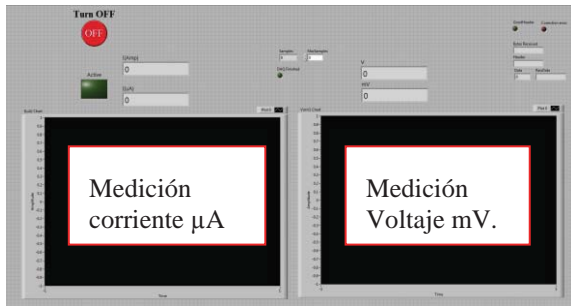


Fig. 27: Pantalla de despliegue de información

**VII.Resultados experimentales del sistema de monitoreo.**

Se realizaron pruebas de funcionamiento del detector GEM, para saber su repuesta normal e identificar los componentes y modos de operación.

En la Fig28 se muestra la conexión estándar de un detector GEM y se indican las ubicaciones del DRIFT, la primera conexión GEM1 y la conexión GEM2, para que se presenten eventos en los detectores, es necesario según lo indicado colocar una pastilla radioactiva que fomentara la aparición de los eventos de detección de partículas cósmicas y electrones, por otra parte incrementara la posibilidad de descargas en el GEM.

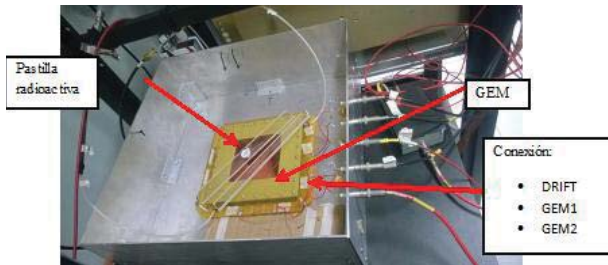


Fig. 28: GEM del Instituto de Ciencias Nucleares

En la Fig. 29 se muestra la conexión al alto voltaje del detector GEM, se puede observar las diferentes capas del detector



Fig. 29: Alimentación de voltajes GEM.

Cabe mencionar que el detector en que se realizan las pruebas cuenta con dos GEM y están conectados según la tabla1:

Nivel del GEM	Descripción	Voltaje aplicado
DRIFT	Capa superior	-1675 V
Top1	Capa superior GEM1	-1650 V
Bottom1	Capa inferior GEM1	-900 V
Top2	Capa superior GEM2	-880 V
Bottom2	Capa inferior GEM2	0 V

Tabla1: Voltajes aplicados en el detector.

Para la conexión del sistema de monitoreo se considera la fabricación de una jaula de Faraday con la finalidad que al sistema se le induzca la menor cantidad de ruido. Se conecta el sistema en el DRIFT del detector. En la Fig. 30 se muestra la conexión del detector GEM, mostrando el cable para el alto voltaje, además se muestra la forma de conexión del sistema de monitoreo.

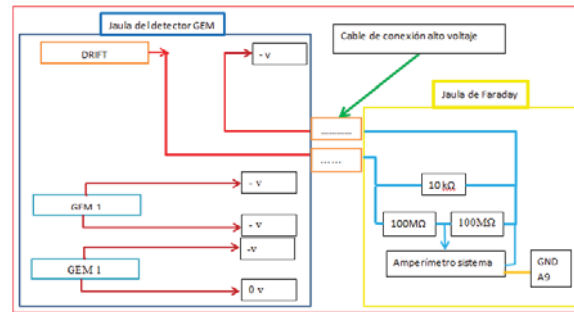


Fig. 30: Diagrama de conexión del sistema de monitoreo y el detector GEM.

En la fig. 31 se muestra la caja de Faraday que se fabricó para poder conectar el sistema al GEM, para lo cual cuenta con conectores de alto voltaje, cabe mencionar que la tarjeta generadora de señal para el transformador se encuentra afuera.

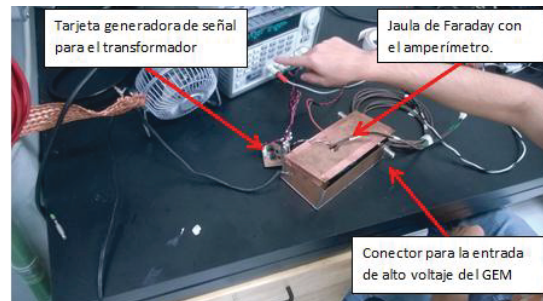


Fig. 31: Tarjeta de monitoreo de corriente en la jaula de Faraday fabricada.

En la Fig. 32 se puede observar la conexión del detector con el sistema de monitoreo.

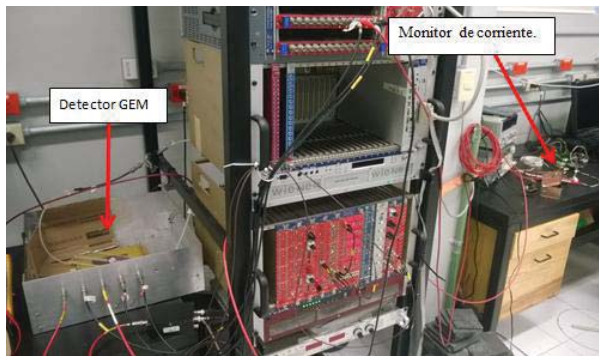


Fig. 32: Conexión del detector con el sistema de monitoreo.

Se realizaron las pruebas del sistema de monitoreo con el detector del ICN. Se observó que el sistema no interfiere con la operación normal del detector, ya que la intensidad de los eventos que mide el detector siguen apareciendo con normalidad y además no induce ruido en la señal de salida.

Con las pruebas que se realizaron en el detector se determinó el ruido intrínseco del sistema de monitoreo, y se observó un ruido de la tarjeta de 60nA y en pruebas posteriores se logró bajo condiciones alcanzar un ruido de 20nA, esta prueba se muestra en la Fig. 33.



Fig. 33: Respuesta del sistema de monitoreo sin energizar el detector.

En la Fig. 34 se muestra la respuesta del sistema de monitoreo. Se logró observar la respuesta del sistema detectando los eventos de descarga del detector GEM de acuerdo a lo esperado.

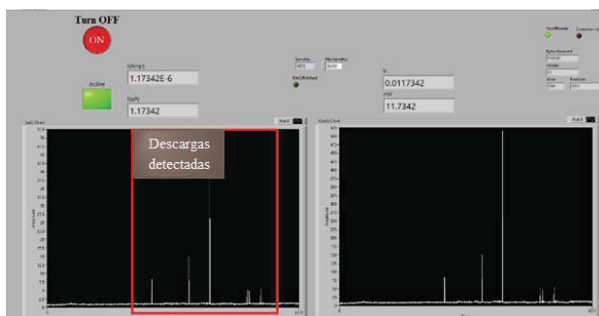


Fig. 34: Respuesta del sistema de monitoreo.

## VIII. Trabajos futuros y conclusiones.

Se instrumentó e implementó el sistema de monitoreo de corriente del detector GEM que se encontrara en la TPC del experimento ALICE del CERN, se observa que no interfiere en la operación normal del detector, además se obtuvieron las mediciones de consumo de corriente del detector en las descargas que se provocaron por los eventos de detección de partículas. Se logró que el sistema tuviera un ruido intrínseco de 20nA en las mediciones.

Como mejoras al sistema de monitoreo se considera implementar un enlace óptico para la transmitir la información, reemplazando los cables LEMO implementados actualmente, esto implica la instrumentación necesaria para el enlace, la implementación de un segundo FPGA. Además la implementación de un ADC de 18 Bits con la finalidad de aumentar la resolución de las mediciones.

## Referencias

- [1]. <http://gdd.web.cern.ch/GDD>, CERN web page, fecha de consulta:13-01-2013.
- [2]. <http://home.cern/about/experiments/alice>, CERN web page, Fecha de consulta: 20-01.2013.
- [3]. JOACHIM BAECHLER , ALICE collaboration, “ The ALICE TPC”, CERN nota internal,paguinas, año.
- [4]. ALICE TPC collaboration, “Technical Design Report of the ALICE TPC Upgrade”, October 22, 2013.
- [5]. <http://www.farnell.com/datasheets/1578598.pdf>, Datasheet component, fecha de consulta: 01-04-2015.
- [6]. <http://www.analog.com/media/en/technical-documentation/data-sheets/AD676.pdf>, Datasheet component, fecha de consulta: 10-06-2015.
- [7]. [http://pdf.datasheetcatalog.com/datasheets2/24/244790\\_1.pdf](http://pdf.datasheetcatalog.com/datasheets2/24/244790_1.pdf), Datasheet component, fecha de consulta: 04-05-2015.
- [8]. [http://host.uniroma3.it/laboratori/escher/res/ES\\_P\\_III/Data%20Sheet/TL081.pdf](http://host.uniroma3.it/laboratori/escher/res/ES_P_III/Data%20Sheet/TL081.pdf), Datasheet component, fecha de consulta: 20-01-2015.

# Diseño y construcción de un sistema de adquisición para el experimento de caminatas ópticas cuánticas con aplicación en cómputo cuántico

A. León, M. A. D. Vargas, S. Vergara, F. Reyes

Facultad de Ciencias de la Electrónica (FCE), Grupo de robótica, Maestría en Ciencias de la Electrónica Opción en Automatización, Benemérita Universidad Autónoma de Puebla (BUAP). Av San Claudio y 18 Sur S/N, C.U., Edificio 109A, C.P. 72570, Puebla, Pue., México.

V. Velazquez

Facultad de Ciencias, Universidad Nacional Autónoma de México (FC-UNAM), Av. Universidad 3000, Circuito Exterior S/N, Delegación Coyoacán, C.P. 04510, Ciudad Universitaria, D.F., México.

**Abstract**—En el siguiente trabajo se presenta el estudio de un interferómetro generalizado de Mach Zehnder, para una red de 2x2 divisores de haz y la implementación del arreglo experimental. Se realizó un estudio de los fenómenos de los sistemas con enredamiento cuántico y se implementó el experimento para demostrar que los sistemas con enredamiento cuántico pueden violar la desigualdad de Bell, también se realizó el análisis matemático para conocer la probabilidad del estado de un fotón pasando a través de la red de divisores de haz que conforman el arreglo experimental, la implementación física del experimento para comparar la información obtenida experimentalmente con los resultados del análisis matemático. Finalmente se diseñó una tarjeta de adquisición con comunicación WIFI, la cual se utilizará para adquirir las señales que son el resultado de la detección de fotones individuales, además se realizará el procesamiento de la información obtenida por el conteo y las coincidencias entre fotones.

This paper presents the study of generalized Mach Zehnder interferometer, for a network of beam splitters 2x2 and implementation of the experimental setup. A study of the phenomena of quantum entanglement systems was performed and the experiment was implemented to demonstrate that quantum entanglement systems may violate Bell's inequality, mathematical analysis was also performed to determine the probability of the state of a photon passing through the network of beam splitters that make up the experimental arrangement, the physical implementation of the experiment to compare the data obtained experimentally with the results of mathematical analysis. Finally an acquisition card with WIFI communication was designed, which will be used to acquire signals that are the result of the detection of individual photons, also processing the information obtained by counting and overlap between photons are made.

**Index Terms**—Mach Zehnder, desigualdad de Bell, WIFI.

## I. INTRODUCCIÓN

LA manipulación de fotones y átomos individuales representa un avance tecnológico con muchas promesas de aplicaciones en comunicaciones e información. La base de esta tecnología es el aprovechamiento de la información codificada en las componentes cuánticas de estos entes individuales. Así, fenómenos como interferencia cuántica y enredamiento cuántico son conceptos que no tienen una contraparte clásica, por ello es muy importante desarrollar tecnología que permita la automatización de estos.

El objetivo de este proyecto es acelerar el estudio del comportamiento de fotones cuánticos y mejorar con ayuda de las nuevas tecnologías la investigación, estudiando la interferencia cuántica de un fotón consigo mismo y el enredamiento cuántico, además para su análisis se llevó a cabo la implementación de los experimentos ópticos que dan validez a estos principios cuánticos.

Para este experimento óptico fue necesario contar con una fuente de fotones individuales para asegurar que tengamos un estado con un solo paquete de energía indivisible, y así poder trabajar con los fenómenos cuánticos [1]. La fuente que se utiliza para la creación de fotones individuales está diseñada con materiales de respuesta óptica no lineal. En el bombeo de un cristal no lineal con un láser de alta potencia puede producir pares de fotones individuales, en este proceso el fotón llamado de bombeo atraviesa un material o cristal, de respuesta óptica no lineal, con un eje óptico, y espontáneamente decae en dos fotones producto [2].

Si los fotones producto tienen la misma polarización, se dice que se trata de un proceso de Conversión Espontánea Paramétrica Descendente (SPDC, Spontaneous Parametric Down-Conversion). SPDC tipo I; si las polarizaciones son ortogonales entre sí, entonces se trata de SPDC tipo II, para este experimento se realiza el proceso de SPDC tipo II, en el cual hay una convención para nombrar a los fotones producto. Al fotón que tiene la misma polarización que el de bombeo se le llama testigo y al otro se le llama señal.

En este proyecto se estudiará la interferencia cuántica de un fotón consigo mismo al cruzar el fotón señal por una red de divisores de haz de dimensión 2x2 basándonos en un Interferómetro Mach-Zehnder (IMZ), se puede observar un patrón de interferencia cuando un fotón cruza la red de divisores, a este patrón de interferencia le llamamos interferencia cuántica. De esta manera reconstruimos experimentalmente las probabilidades, de que el fotón salga por alguno de los 4 puertos de salida disponibles.

Para la detección de los fotones se utilizaron sensores fotodiodos de avalancha (APDs), para el montaje del arreglo experimental los APDs utilizados constan de unos módulos de

conteo de fotones SPCM-AQ4C Perkin Elmer, los cuales por cada detección de un fotón emiten un pulso TTL de 10 ns., que posteriormente será enviado a un circuito electrónico que realiza el conteo de coincidencias, para esto se requiere realizar el diseño de una tarjeta de adquisición con comunicación WiFi, para realizar los conteos individuales y las coincidencias entre fotones.

Anteriormente, en el año de 2005 el Dr. Sergio Vergara y la Dra. Aurora Vargas, de la FCE-BUAP, diseñaron y construyeron una tarjeta de adquisición de datos. Dicha tarjeta PCI, fue la primera versión un contador de fotones individuales y sus coincidencias, utilizada en la FC-UNAM. Dicha tarjeta emplea un FPGA MAX7000AETC144-5, tiene una resolución de 30ns, y cuenta con el protocolo de comunicación PCI[3], en este trabajo se presenta un nuevo contador de fotones individuales y coincidencias.

April 6, 2016

## II. ENREDAMIENTO CUÁNTICO Y PRUEBA DE LA DESIGUALDAD DE BELL

La importancia de la demostración experimental de la violación de las desigualdades de Bell es que refuta este argumento manteniendo, por ahora, a la mecánica cuántica como descripción correcta de la naturaleza.

Experimentalmente, resulta más práctico el uso de parejas de fotones enredados en polarización, en lugar de parejas de partículas con enredamiento en espines [4].

### A. Enredamiento

Si dos fotones correlacionados con polarización opuesta en un estado producto entran a un divisor de haz por el puerto 1. (Figura 1).

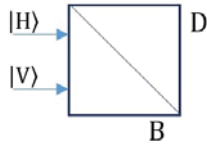


Fig. 1. Dos fotones con polarización perpendicular entrando por el mismo puerto)

$$|\Phi\rangle_e = |H\rangle_1|V\rangle_1 \quad (1)$$

La salida del divisor será

$$|\Psi\rangle_s = \frac{1}{2}[(|H\rangle_D + i|H\rangle_B)[|V\rangle_D + i|V\rangle_B] \quad (2)$$

Reescribiendo:

$$|\Psi\rangle_s = \frac{1}{\sqrt{2}}(i|\psi^+\rangle + |\psi'^-\rangle) \quad (3)$$

donde

$$|\psi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_D|V\rangle_B + |H\rangle_B|V\rangle_D) \quad (4)$$

$$|\psi'^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle_D|V\rangle_D - |H\rangle_B|V\rangle_B) \quad (5)$$

son estados enredados. En una detección en coincidencia, únicamente detectaremos el estado enredado  $|\psi^+\rangle$  [5, 6]. En ambas salidas no es necesario interpretar las direcciones por medio de etiquetas (B y D), debido a que los dos fotones tienen inicialmente el mismo puerto de entrada y por tanto, transmitido y reflejado significa lo mismo a la salida.

Si ahora tenemos la combinación en la que el fotón de polarización horizontal entra por el puerto 1 y el fotón vertical entra por el puerto 2, (Figura 2).

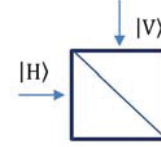


Fig. 2. Dos fotones con polarización perpendicular entrando por diferente puerto

$$|\Phi\rangle_e = |H\rangle_1|V\rangle_1 \quad (6)$$

El resultado en este caso es muy similar que en el caso de los dos fotones entrando por el mismo puerto, con la diferencia de que en el fotón vertical que entre por el puerto 2, debemos cambiar poner las etiquetas de B por D y viceversa.

$$|\Psi\rangle_s = \frac{1}{\sqrt{2}}(i|\psi'^+\rangle + |\psi^-\rangle) \quad (7)$$

donde

$$|\psi^-\rangle = \frac{1}{\sqrt{2}}(|H\rangle_D|V\rangle_B - |H\rangle_B|V\rangle_D) \quad (8)$$

$$|\psi'^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_D|V\rangle_D + |H\rangle_B|V\rangle_B) \quad (9)$$

Este es un estado de Bell que sólo puede ser detectado en coincidencia .

### B. Estado Enredado Inicial

Al introducir un estado enredado  $|\psi^+\rangle$  en el divisor de haz implicaría que a veces una parte entra por el puerto 1 y en ocasiones por el puerto 2, y viceversa, (Figura 3).

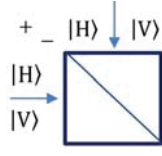


Fig. 3. Dos fotones con polarización perpendicular entrando por diferente puerto

Si el estado de entrada es un estado  $|\psi^+\rangle = \frac{1}{\sqrt{2}}(|H\rangle_1|V\rangle_2 + |V\rangle_1|H\rangle_2)$ , entonces, el estado de salida puede escribirse como:

$$|\Psi\rangle_s = \frac{1}{2\sqrt{2}}[(|H\rangle_D + i|H\rangle_B)(|V\rangle_B + i|V\rangle_D) + (|V\rangle_D + i|V\rangle_B)(|H\rangle_B + i|H\rangle_D)] \quad (10)$$

Desarrollando los productos y posteriormente cancelando los terminos correspondientes, obtenemos:

$$|\Psi\rangle_s = \frac{i}{2\sqrt{2}}(|H\rangle_D|V\rangle_D + |H\rangle_B|V\rangle_B + |V\rangle_D|H\rangle_D + |V\rangle_B|H\rangle_B) \quad (11)$$

Simplificando:

$$|\Psi\rangle_s = \frac{i}{\sqrt{2}}(|H\rangle_D|V\rangle_D + |H\rangle_B|V\rangle_B) \quad (12)$$

Comparando con la ecuación 9

$$|\Psi\rangle_s = i|\psi'^+\rangle \quad (13)$$

Este estado es reversible, es decir, ahora intruduciendo un estado  $|\psi^+\rangle$ , obtendremos el estado  $|\psi^+\rangle$ . Debido a esto se pueden decir lo siguiente:

En el caso de dos fotones idénticos entrando por los dos puertos del divisor de haz, el efecto principal es la interferencia destructiva de la combinación de estados de fotones individuales saliendo por los dos puertos de salida. Sin embargo, cuando los fotones tienen polarización opuesta, el efecto de la interferencia sería impensable si el estado no estuviera enredado, como se muestra en las ecuaciones 3 y 6. Si un estado enredado de fotones con polarización opuesta es introducido al divisor de haz, podemos observar que pueden hacer interferencia.

Ya que el divisor de haz es un elemento óptico pasivo, en base a los resultados, podemos decir que la correlación del estado entrante debe conservarse. Si el estado entrante no tiene enredamiento, entonces el estado saliente tampoco, el estado saliente será una combinación de estados enredados que sumados no tienen enredamiento. Si lo que es introducido

al divisor de haz es un estado enredado, entonces el estado saliente será un estado enredado también.

En cuanto a la teoría de la desigualdad de Bell, según la mecánica cuántica, los valores numéricos que definen las propiedades de un objeto (como su posición, momento o espín) no se encuentran completamente determinados hasta que no se miden. Eso implica que la teoría no predice de manera única el resultado de un experimento, sino solo probabilidades: por ejemplo: al hacer el experimento E y medir la cantidad A, los únicos resultados posibles son  $a_1$ ,  $a_2$  y  $a_3$ , y la probabilidad de obtener cada uno es  $p_1$ ,  $p_2$  y  $p_3$ .

En noviembre de 1964, John Stewart Bell, propuso un experimento que planteaba un montaje experimental con dos partículas entrelazadas y demostraba que, bajo ciertas condiciones, las predicciones de la mecánica cuántica eran incompatibles con de las de cualquier otra teoría basada en variables ocultas locales (es decir, cuyos efectos no pudiesen propagarse de manera instantánea entre puntos distantes del espacio). [7]

### III. INTERFERENCIA CUÁNTICA

En el Interferómetro de Mach Zehnder, un haz de luz coherente es separado en dos rayos de luz por un divisor de haz. Cada haz sigue un camino diferente al del otro (camino óptico) y ambos son recombinados antes de ser detectados. La diferencia en la distancia recorrida por cada haz produce una diferencia de fase entre ellos al momento de hacer reincidir los rayos de luz, la diferencia de fase entre ellos produce un patrón de anillos concéntricos o franjas de luz espaciadas entre ellas producto de las interferencias constructivas y destructivas

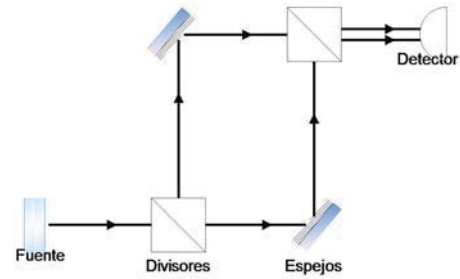


Fig. 4. Interferómetro de Mach-Zehnder

Para analizar la interferencia con fotones individuales, consideremos el arreglo del Interferómetro de Mach Zehnder. Cuando un fotón entra a un divisor de haz, su salida puede escribirse de la siguiente manera:

$$|\psi\rangle = C_T|1\rangle^T + iC_R|1\rangle^R \quad (14)$$

Donde  $|\psi\rangle$  es la función de estado del fotón que cruza el divisor de haz, que es igual a la combinación lineal de los

coeficientes  $C_T$  y  $C_R$ . Donde  $|1\rangle^T$  se refiere al estado del fotón transmitido y  $i|1\rangle^R$  al estado del fotón reflejado. La ecuación 14. Nos indica que el estado  $|\psi\rangle$  va a ser igual una amplitud de probabilidad  $C_T$  de que el fotón sea transmitido más una amplitud de probabilidad  $C_R$  de que el fotón sea reflejado, si el divisor de haz es 50:50, la probabilidad de ser transmitido y de ser reflejado es 50%. Véase la figura 5.

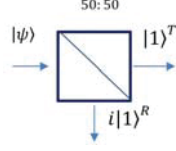


Fig. 5. Combinación de posibilidad del estado de un fotón pasando a través de un divisor de haz.

Si  $|C_T|^2$  y  $|C_R|^2$  se refiere a la probabilidad del fotón transmitido y del fotón reflejado respectivamente en la salida del divisor, dado que la probabilidad es la misma en ambas salidas (50:50) tenemos que:

$$|C_T|^2 = |C_R|^2 \quad (15)$$

De acuerdo al principio de conservación de la probabilidad:

$$|C_T|^2 + |C_R|^2 = 1 \quad (16)$$

Sustituyendo la ecuación 15 en la ecuación 16 y despejando:

$$C_T = \frac{\pm 1}{\sqrt{2}} \quad (17)$$

Sustituyendo en la ecuación 14, se puede reescribir como:

$$|\psi\rangle = \frac{1}{\sqrt{2}}|1\rangle^T + i\frac{1}{\sqrt{2}}|1\rangle^R \quad (18)$$

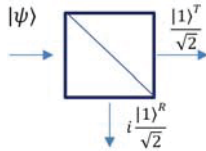


Fig. 6. Combinación de probabilidad del estado de un fotón pasando a través de un divisor de haz.

1) *Probabilidad para dos divisores en serie horizontal:* Sea  $|\psi_1\rangle$  la función de estado para el segundo divisor tenemos:

$$|\psi_1\rangle = C_{T1}\frac{|1\rangle^T}{\sqrt{2}} + iC_{R1}\frac{|1\rangle^T}{\sqrt{2}} \quad (19)$$

Realizando la metodología para hallar la función de estado a la salida del primer divisor de haz, tenemos que para este caso la función de salida será:

$$|\psi_1\rangle = \frac{|1\rangle^T}{\sqrt{2}} + i\frac{|1\rangle^T}{\sqrt{2}} \quad (20)$$

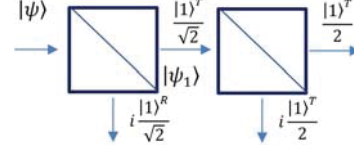


Fig. 7. Combinación de probabilidad del estado de un fotón pasando a través de divisores de haz en serie horizontal.

2) *Probabilidad para dos divisores en serie vertical:* Sea  $|\psi_2\rangle$  la función de estado para el segundo divisor en vertical:

$$|\psi_2\rangle = C_{T2}i\frac{|1\rangle^T}{\sqrt{2}} + iC_{R2}i\frac{|1\rangle^R}{\sqrt{2}} \quad (21)$$

Finalmente obtenemos función de estado de la salida:

$$|\psi_2\rangle = i\frac{|1\rangle^T}{2} - \frac{|1\rangle^T}{2} \quad (22)$$

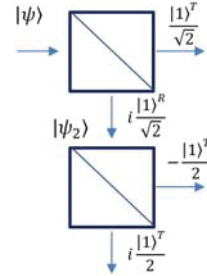


Fig. 8. Combinación de probabilidad del estado de un fotón pasando a través de divisores de haz en serie vertical.

3) *Arreglo de 2 x 2 divisores de haz:* Tenemos que la función de estado para  $|\psi_3\rangle$  es la suma de cada una de las componentes resultantes de las funciones de estado  $|\psi_{1R}\rangle$  y  $|\psi_{2R}\rangle$

Sumando las componentes que salen por la derecha del divisor:

$$-\frac{|1\rangle^T}{2\sqrt{2}} - \frac{|1\rangle^R}{2\sqrt{2}} = -\frac{|1\rangle}{\sqrt{2}} \quad (23)$$

Sumando las componentes que salen por abajo del divisor:



$$i \frac{|1\rangle^T}{2\sqrt{2}} - i \frac{|1\rangle^R}{2\sqrt{2}} = 0 \quad (24)$$

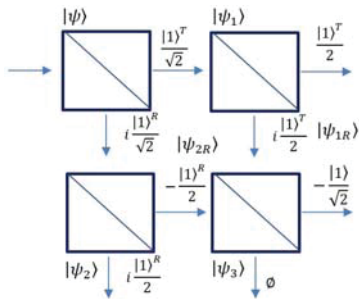


Fig. 9. Combinación de probabilidad del estado de un fotón pasando a través de una red de  $2 \times 2$  divisores de haz

Dado el principio de conservación de la probabilidad podemos comprobar que se cumple para las salidas de este arreglo

$$\left| \frac{1}{\sqrt{2}} \right|^2 + \left| -\frac{1}{\sqrt{2}} \right|^2 + |0|^2 + \left| \frac{1}{\sqrt{2}} \right|^2 = 1 \quad (25)$$

#### IV. DISEÑO DE UN SISTEMA DE ADQUISICIÓN PARA AUTOMATIZAR EL EXPERIMENTO

Para el arreglo opto-electrónico de este experimento se requiere una fuente de fotones individuales, divisores de haz, detectores y un sistema de conteo, la Figura 10 muestra un esquema general de este experimento.

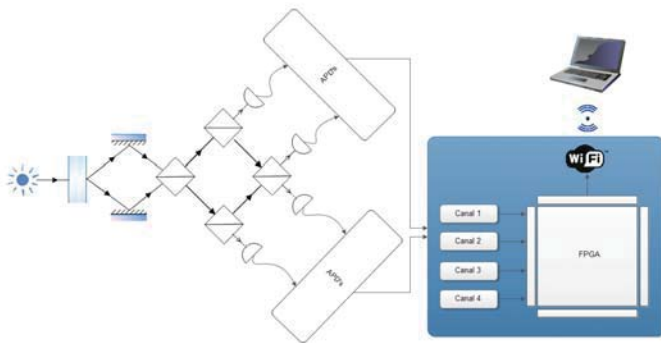


Fig. 10. Esquema General del arreglo optoelectrónico

##### A. Hardware

El hardware, consiste en una tarjeta modular, contiene una tarjeta madre y una hija.

1) *Tarjeta hija*: La tarjeta hija, se conecta a la tarjeta madre, es un diseño de 4 capas que ya ha sido diseñada previamente en la FCE-BUAP, esta posee un FPGA Cyclone III de Altera (EP3C10F256C6), y también se puede programar y utilizar conectándola a una fuente de alimentación externa. (Figura 7b).

*Parte superior*: Consta de: 1. Un FPGA Cyclone III con matrícula EP3C10F256C6 de Altera; 2. Un puerto de programación; 3. Un puerto JTAG; 4. Un conector para alimentación externa; 5. Un botón de reinicio; Conectores header macho de 50 pines.



Fig. 11. Parte superior de la tarjeta hija

*Parte inferior*: Consta de: 1. Una memoria EEPROM (EPC2); 2. Un convertidor DC-DC de 3.3V a 1.2V; 3. Un Oscilador de cristal; 4. Resistencias y capacitores.

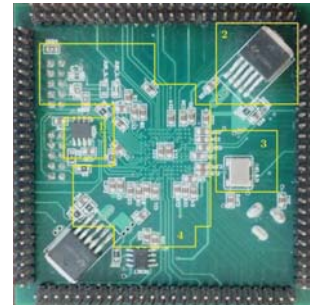


Fig. 12. Parte inferior de la tarjeta hija

2) *Tarjeta madre*: La tarjeta madre consta de cuatro capas de cobre y tres de material dieléctrico, Las pistas para las señales se localizan en la capa superior e inferior; El plano de alimentación y tierra se encuentran en las capas intermedias.

Los componentes para la tarjeta madre son: 1. Nueve conectores LEMO; 2. Cuatro conectores Header de 50 pines cada uno, para conectar la tarjeta hija; 2. Dos conectores Header para el módulo Wi-Fi; 3. Un conector Header de 25 pines para propósito general; 4. Resistencias para divisores de voltaje y capacitores.

La salidas de los sensores APDs se conectan a la tarjeta madre mediante conectores LEMO, Debido a que la señal de cada canal es un pulso TTL de 4.5 Volts y el FPGA trabaja con un nivel de voltaje 3.3 Volts, fue necesario implementar un divisor de voltaje a la salida de cada conector. Al primer banco

del FPGA se le ha asignado los nueve canales de entrada, el banco dos se ha destinado para conexiones de proposito general, y en el banco tres para las señales para el protocolo de comunicación UART. Para la conexión al FPGA de la interfaz Wi-Fi se realizó un cableado de dos líneas ya que el modulo RN-XV cuenta con el protocolo serial UART. El diagrama a bloques del circuito se muestra en la figura 13., además se colocaron capacitores de  $0.1\mu\text{F}$ . conectados de Vcc a GND para reducir la interferencia electromagnética.

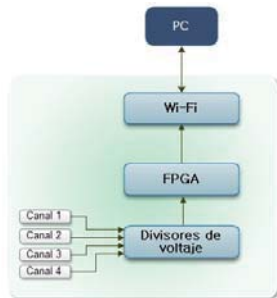


Fig. 13. Diagrama a bloques del circuito

Posteriormente se diseña la tarjeta de circuito impreso (PCB, Printed Circuit Board) del hardware. Para este propósito, se utilizó el software de diseño Altium Designer v14.3.9 [8]

En esta capa se muestran las pistas de conexión entre los canales de entrada, divisores de voltaje y las pistas entre la tarjeta madre y el módulo Wi Fi.

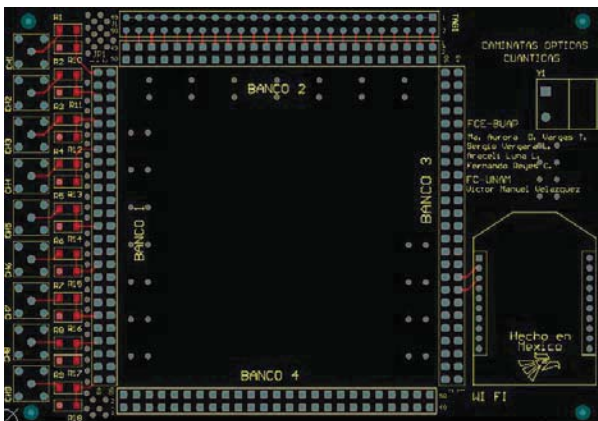


Fig. 14. Top Layer o Capa superior

3) *Capa de tierra:* En la Figura se muestra la capa de GND o capa de tierra

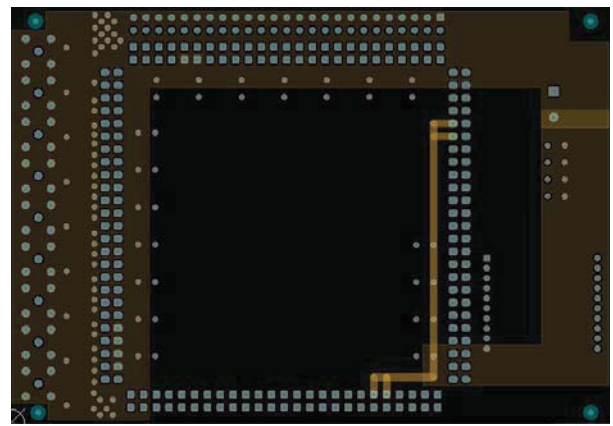


Fig. 15. Capa de tierra

4) *Capa de alimentación:* En la Figura se muestra la de alimentación o VCC.

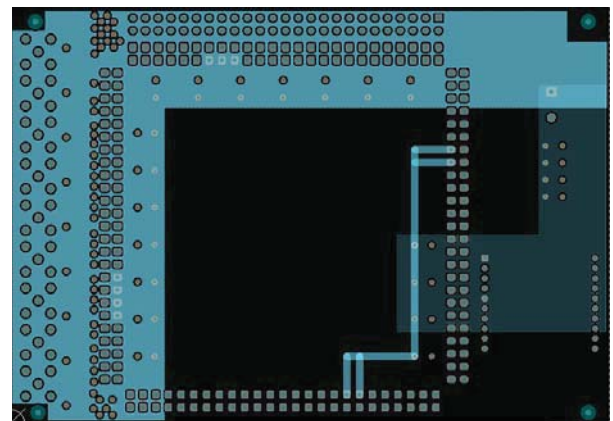


Fig. 16. Capa de alimentacion o VCC.

5) *Capa inferior:* En esta capa se montan los capacitores de reducir la interferencia electromagnica que se genera entre pistas.

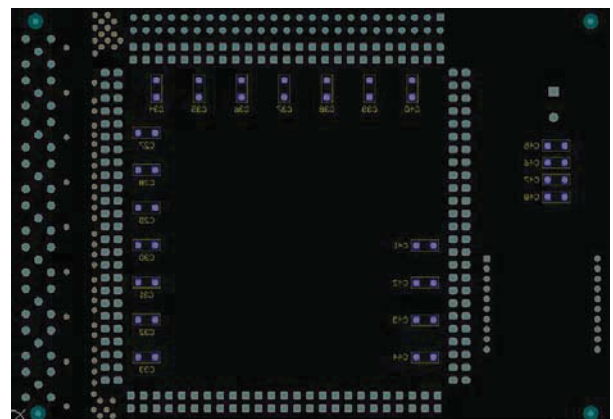


Fig. 17. Capa inferior o Bottom Layer

**B. Firmware**

El firmware a desarrollar se implementará en un FPGA Cyclone III de Altera (EP3C10F256C6) el cual se encuentra montado en la tarjeta hija mencionada anteriormente. La arquitectura de este FPGA consiste de hasta 120,000 elementos lógicos dispuestos verticalmente, 4 Mbits de memoria integrada dispuestos como bloques de 9-Kbit (M9K) y 200 multiplicadores integrados 18x18 bits.

En la Figura 18 se muestra un diagrama a bloques del firmware propuesto.

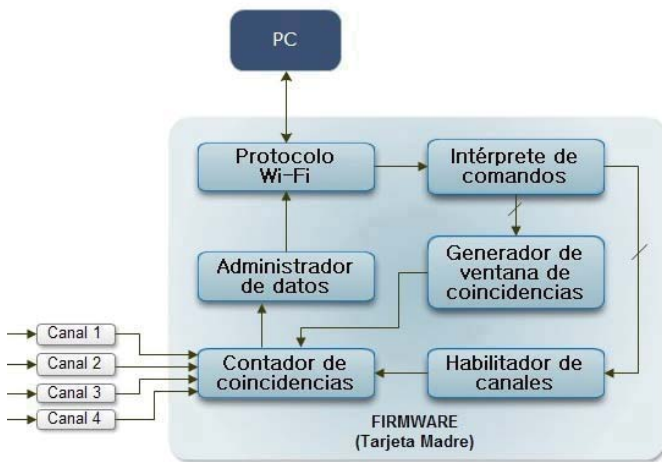


Fig. 18. Diagrama a bloques del Firmware

El Protocolo Wi-Fi, establece la comunicación entre el firmware y software. El intérprete de comandos decodifica las órdenes del software y las envía al habilitador de canales y al generador de ventanas de coincidencias. El habilitador de canales selecciona los canales que estarán disponibles y el generador de ventana de coincidencias genera las señales de coincidencia entre señales. El contador de coincidencias cuenta las coincidencias existentes entre las señales y envía esta información al administrador de datos, quien las transmite al software para realizar su procesamiento.

**C. Software**

El software que se propone esta implementado en LabView. Este se ejecuta en una PC y permite seleccionar los canales de entrada, introducir el tiempo de la ventana de coincidencias y de prueba, almacenar los datos del experimento, introducir el tiempo de la prueba, y finalmente mostrar los resultados

En la Figura 19 se muestra una interfaz gráfica del software propuesto.



Fig. 19. Interfaz gráfica del software propuesto

También la interfaz gráfica del software permitiera al usuario poder ingresar los 16 ángulos de las correlaciones para poder obtener el parámetro de la desigualdad de Bell. En la Figura 20 se muestra una interfaz gráfica.



Fig. 20. Interfaz gráfica para el parámetro de la desigualdad de Bell

Finalmente la interfaz gráfica de la figura 21 nos muestra el resultado del conteo de coincidencias de fotones.

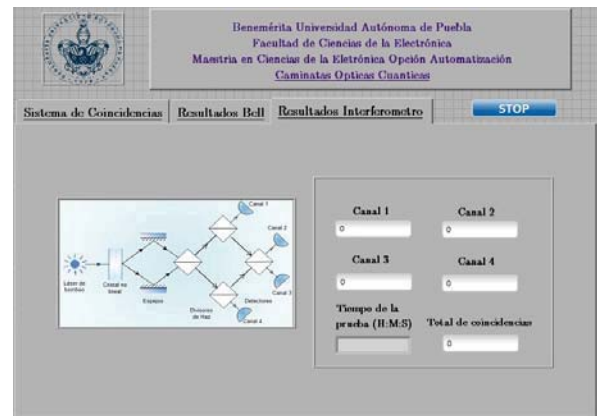


Fig. 21. Interfaz gráfica para el parámetro del resultado de conteo de coincidencias

## V. RESULTADOS EXPERIMENTALES

Para el experimento que nos permite obtener de las mediciones que comprueban la violación de la desigualdad de Bell se utilizó el Láser de bombeo de intensidad variable, azul violeta CL-2005 sintonizable, de 5 a 100 mW, de 410 nm (Crystal Laser). La parte medular de este experimento es la fuente de fotones individuales. Para crearlos, utilizamos un Cristal BBO Tipo II (Newlight Photonics Inc.).

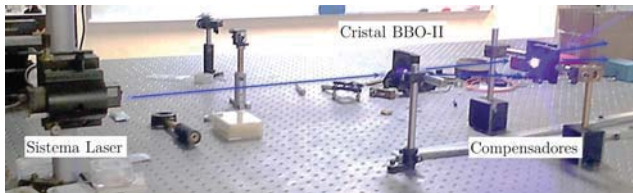


Fig. 22. Montaje del experimento (Cruce de los fotones a través de los compensadores)

Para estudiar correlaciones de polarización en la pareja de fotones se realiza lo siguiente:

- Modificar la polarización de los fotones señal y testigo antes de ser detectados.
- Permitir el paso únicamente de fotones con una cierta polarización lineal (polarizador o desfases).

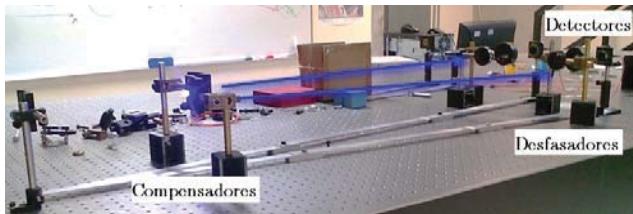


Fig. 23. Montaje del experimento (Cruce de los fotones a través de los polarizadores o desfases)

Al realizar 16 mediciones de probabilidades de detección conjunta de fotones o tasas de coincidencias. Se pudo probar que la desigualdad de es equivalente a  $-2 \leq S \leq 2$ . En este experimento se obtuvo  $S = -2.32$ , lo cual implica una clara violación a la desigualdad de Bell. La demostración experimental de que sistemas con enredamiento cuántico pueden violar la desigualdad de Bell es de enorme importancia ya que puede considerarse como una refutación de las teorías de variables ocultas, sostenedoras de los principios de realismo y de localidad. La evidencia experimental indica la validez de la mecánica cuántica como descripción correcta de la naturaleza a muy pequeña escala.

El experimento se realizó con una red de divisores de  $2 \times 2$  (Figura 24), donde el láser violeta cruza el divisor BBO-II para poder obtener fotones individuales (un testigo y un señal), el fotón testigo realiza su camino óptico a través del arreglo de divisores de  $2 \times 2$ .

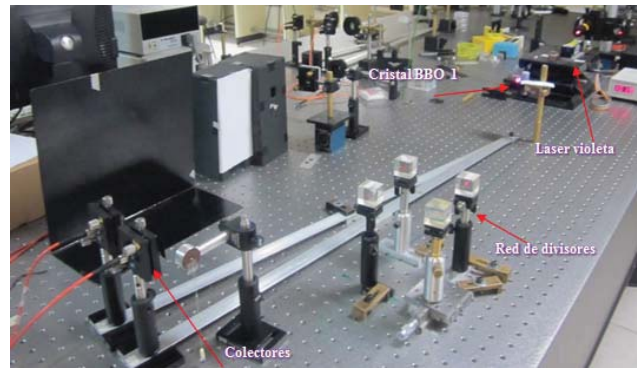


Fig. 24. Montaje del arreglo experimental de  $2 \times 2$  divisores de haz.

Por otra parte para obtener el patrón de interferencia a la salida del divisor (Figura 25), en el arreglo experimental la probabilidad cuántica del encontrar un fotón de un lado de la cara del cristal es nula y en la otra cara es  $\frac{1}{2}$  en el caso de que la probabilidad en el cristal sea exactamente de 50:50, por ello se eligió realizar el conteo de coincidencias del lado donde obtendremos los máximos de las franjas de interferencia.

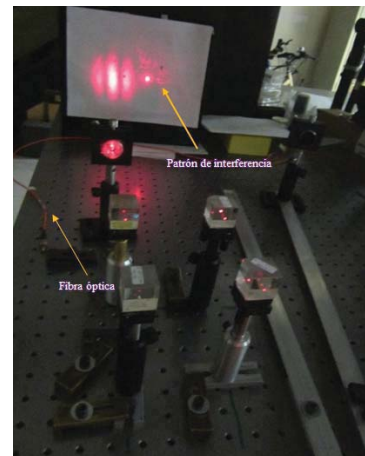


Fig. 25. Patrón de interferencia de los máximos

Al modificar el camino cuántico del fotón desplazando linealmente el divisor de haz con ayuda del piezoeléctrico también se modifica el patrón de interferencia y a su vez las probabilidades cuánticas en cada una de las salidas de los divisores de haz se alteran. Al realizar el conteo de coincidencias durante el desplazamiento lineal se obtiene la ubicación espacial del fotón (Figure 26), Al observar el conteo de coincidencias en cada salida de los divisores, podemos comprobar que cumple con el análisis matemático que se realizó para encontrar la probabilidad cuánticas de hallar el fotón en cada una de las salidas de los canales.

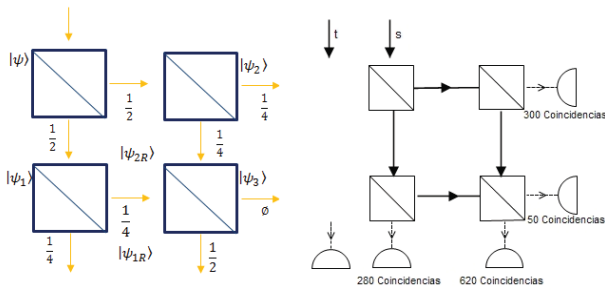


Fig. 26. Patrón de interferencia de los máximos

## VI. TRABAJO A FUTURO

Las pruebas preliminares que se realizaron en el laboratorio de Óptica Avanzada de la FC-UNAM sobre una arreglo opto-electrónico mostraron resultados satisfactorios, además de que la adquisición de los datos se realizó de forma manual. Debido a que las características de la tarjeta de adquisición requiere una fabricación especializada, esta se encuentra en el proceso de fabricación en el extranjero, posteriormente para la implementación del experimento se pretende realizar con un arreglo opto-electrónico fabricado dentro de un chip óptico, y en cuanto se cuente con la tarjeta se realizarán las pruebas experimentales y se comparará la eficiencia del sistema automatizado con los resultados obtenidos de forma manual.

## REFERENCES

- [1] J. López A., “Automatización del Experimento de correlación de pares de fotones de conversión paramétrica descendente,” Master’s thesis, Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, 2013.
- [2] J. J. López-Durán, “Caracterización de un cristal BBO tipo II para su uso como fuente de estados entrelazados,” *Cinvestav*, p. 8.
- [3] P. G. T. G. F. A. L. I. R. F. Vergara S., Vargas M.A. and V. L., “Characterization of the ACORDE scintillator counters using a PCI electronic card,” *Revista Mexicana de Física*, vol. 53, pp. 120–125, 2007.
- [4] E. B. V. V. E. L.-M. y. M. G. Luis Edgar Vicent, Alfred B. U’Ren, “Fotones enredados y desigualdades de Bell: explorando la no-localidad,” *Cinvestav*, p. 11, enero-marzo 2008.
- [5] Y. Shih and C. Alley, “New type of Einstein-Podolsky-Rosen-Bohm experiment using pairs of light quanta produced by optical parametric down conversion,” *Phys. Rev.*, 1988.
- [6] Z. U. Ou and L. Mandel, “Violation of Bell’s Inequality and Classical Probability in a Two-Photon Correlation Experiment,” *Phys. Rev.*, oct. 1988.
- [7] I. y ciencia, “El teorema de bell cumple 50 años,” web; accedido el 03-02-2016. [Online]. Available: <http://www.investigacionyciencia.es/noticias/el-teorema-de-bell-cumple-50-aos-12686>

- [8] Microensamble, “Precauciones que se deben tener en cuenta al generar planos de tierra o vcc,” 23 de octubre de 2014, web; accedido el 03-02-2016. [Online]. Available: <http://microensamble.com/blog/cuidados-que-se-deben-tener-con-los-planos-de-tierra/>

# Sistema de Sensorización no intrusiva en Entornos con Inteligencia Ambiental, a partir de casos de emergencia.

Castañeda Delgado, Jaime  
División de Electromecánica  
Instituto Tecnológico Superior Zacatecas Norte  
Río Grande, Zacatecas  
[jaime.castaneda@itszn.edu.mx](mailto:jaime.castaneda@itszn.edu.mx)

Molina Wong, María del Refugio  
División de Electromecánica  
Instituto Tecnológico Superior Zacatecas Norte  
Río Grande, Zacatecas  
[maria.molina@itszn.edu.mx](mailto:maria.molina@itszn.edu.mx)

Salas Torres, José María  
División de Informática y Sistemas  
Instituto Tecnológico Superior Zacatecas Norte  
Río Grande, Zacatecas  
[jose.salas@itszn.edu.mx](mailto:jose.salas@itszn.edu.mx)

Delgado Guillén, Lorenzo Antonio  
División de Informática y Sistemas  
Instituto Tecnológico Superior Zacatecas Norte  
Río Grande, Zacatecas  
[lorenzo.delgado@itszn.edu.mx](mailto:lorenzo.delgado@itszn.edu.mx)

Esquivel Salas, Abraham  
División de Informática y Sistemas  
Instituto Tecnológico Superior Zacatecas Norte  
Río Grande, Zacatecas  
[abraham.esquivel@itszn.edu.mx](mailto:abraham.esquivel@itszn.edu.mx)

**Resumen**— Los Entornos con Inteligencia Ambiental integran muchos y variados dispositivos de cómputo (ubiquitous computing), dispersos e integrados de tal manera que pasan desapercibidos a los usuarios, ofreciendo servicios personalizados sensibles al contexto. En este artículo, se describe un proyecto de investigación que pretende poner los Entornos con Inteligencia Ambiental al servicio de personas vulnerables (tercera edad), motivados por el deseo de contribuir a mantener su independencia y calidad de vida en ese segmento de población que cada vez tiende a ser más cuantioso y menos atendido. El estudio fue realizado en un Entorno Activo real que sirve como laboratorio de pruebas para líneas de investigación afines a la Inteligencia Ambiental.

**Keywords**— *Ubiquitous Computing; Pervasive Computing; elderly people*

## I. INTRODUCCIÓN

Con una tendencia creciente de la población de la tercera edad, tendencia que se mantiene en México según las proyecciones de la CONAPO, y una tendencia contraria y decreciente de la población económicamente activa, se justifica el empleo de la tecnología, en específico de los Entornos con Inteligencia Ambiental con el objetivo de mejorar la calidad de vida de este sector de población al que referenciamos como personas vulnerables. De hecho, una de las principales

preocupaciones para las Instituciones encargadas de proporcionar Salud, radica en aliviar la demanda de este sector de personas vulnerables [1], para una población que por factores económicos, culturales, además de la falta de alternativas tales como asilos u otras dependencias que puedan cuidar de este sector de personas vulnerables, están obligados a pasar su vejez solos, en casa. En este sentido, los Entornos con Inteligencia Ambiental ofrecen soluciones que van desde el monitoreo de signos vitales, seguimiento de la localización, seguridad, entre otros servicios que pueden ayudar a mejorar la calidad de vida de este sector de personas vulnerables. Los Entornos con Inteligencia Ambiental pueden detectar cambios en los hábitos de las personas que pueden ser preocupantes [2], lo que puede funcionar como un sistema de alertas para las personas a cargo de cuidar a estas personas vulnerables. Los Entornos con Inteligencia Ambiental puede dar el soporte necesario para que las personas de la tercera edad, puedan permanecer en casa y recibir la atención y cuidados necesarios de una manera mas simple y cada vez menos intrusiva.

## II. MOTIVACIÓN

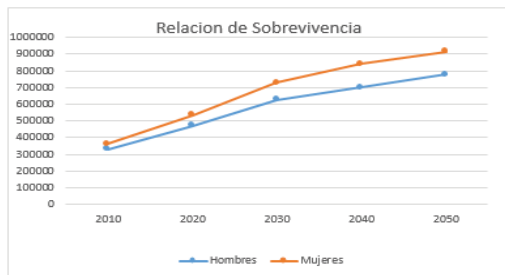
La población de Estados Unidos de América, según el censo de 1996 reportó una población de cerca de 44 millones de adultos con una edad promedio de 60 años, y se prevee que lleguen a 82 millones en 2025.

La importancia de esta propuesta es evidente, cuando se conoce que actualmente en México solo el 2% de la población de personas mayores de 60 años cuenta con algún tipo de alojamiento institucional (casa hogar, asilo etc.). El porcentaje restante, mantiene su estilo de vida, dependiendo de los cuidados de su familia, y por ende, son candidatos a usar la Inteligencia Ambiental como método para mantener su bienestar.

En México, actualmente la población de personas mayores a 60 años es de 12% y se prevé que para las 4 siguientes décadas (2050), la población de personas de la tercera edad tenga un crecimiento de más del doble de su población actual, según datos de la CANAPO 2015. (consejo Nacional de Población) como se puede ver en la figura 1.

**Estimaciones y Proyecciones de población de la República Mexicana según CANAPO periodo año, 2010-2050 (Consejo Nacional de Población 2015).**

Grafica de Indicadores Demográficos, de Personas de 60 años en etapa no productiva.



Lo que se observa de estos indicadores es que la población de personas de la tercera edad tiene un crecimiento en un 114.9% del periodo del año de 2010 al 2050, es decir duplica su población.

Fig. 1. Proyección al año 2050 de personas de 60 años en etapa no productiva, según la CANAPO.

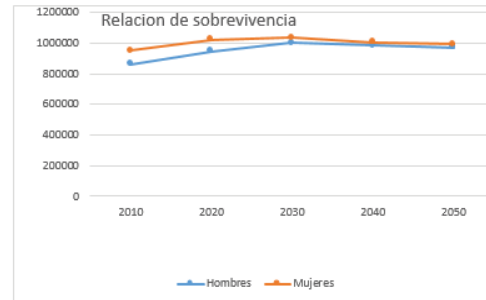
Por otra parte, se prevé que la población económicamente activa que de algún modo puede solventar la infraestructura de salud, decrecerá en un 2.8% en las 4 primeras décadas según datos de la CANAPO (véase la figura 2).

Por razones de elección personal y económica, muchas personas vulnerables vivirán solas en un futuro. Y un gran número ya lo hace.

Un Entorno con Inteligencia Ambiental, con aplicaciones sensibles al contexto pueden aportar beneficios para el cuidado de la salud en personas mayores. Las personas mayores son particularmente propensos a accidentes y caídas en el hogar, quedando heridos por largos periodos de tiempo. También su situación puede hacerlos vulnerables de la delincuencia lo que los lleva a tener una baja calidad de vida a consecuencia de ese miedo, o pueden tener problemas cognitivos, unos de los más típicos es la falta de memoria que puede causar problemas de salud como la gente que olvida tomar sus medicamentos habituales.

Los usuarios también pueden enfrentarse a problemas auditivos y limitaciones motoras, lo que puede provocar accidentes (incendios, fugas de gas etc.). Estos son temas críticos que hablan de la importancia de este tipo de sistemas.

Grafica de Indicadores Demográficos, de Personas de 30 años en etapa productiva.



Para estos indicadores se concluye que para personas económicamente activas para el periodo del año 2010-2030 la población tiene un incremento del 13.9% después para el periodo del año 2030-2050 la población empieza a decrecimiento en un 2.8%.

Fig. 2. Proyección al año 2030 de personas económicamente activas, según la CANAPO.

III. ESTADO DEL ARTE

Mozer Adaptive House, proyecto de la Universidad de Colorado (<http://www.cs.Colorado.edu/~mozer/house/>), consiste de un Entorno capaz de aprender los hábitos de sus habitantes a partir de redes neuronales, y adaptar sus propiedades físicas al usuario a través de una serie de sensores y actuadores diseminados por toda la estructura. Un aspecto relevante es el proceso de aprendizaje acerca de la conservación de los recursos, ahorrando luz y calor, a menos que el usuario determine lo contrario.

En cuanto a su infraestructura, emplea una combinación de sensado remoto y tecnología X10, por lo que algunos de sus sensores están conectados directamente a la red eléctrica.

The Aware Home [3], es un Entorno Activo capaz de sentir la actividad física de los habitantes de un hogar. Su aportación es relevante e innovadora, por una parte, desarrollaron un portaretratos digital, que desde remoto y empleando el marco como medio de visualización de resultados, muestra figuras en movimiento, y la velocidad con que se muestran las figuras está relacionada con la actividad dentro del Entorno Activo. Esto permite que las personas responsables de los usuarios del Entorno, puedan monitorear de manera poco intrusiva su estado de salud.

A manera de crítica, en ambos proyectos los Entornos Activos carecen de la generación de alertas que sirvan como retroalimentación acerca de la información que otros están viendo de los usuarios del Entorno Activo.

Identify applicable sponsor/s here. If no sponsors, delete this text box (sponsors).

#### IV. OBJETIVOS

Diseñar e implementar un sistema de sensorización que permita una vigilancia poco intrusiva para el sector de personas vulnerables, compuesto principalmente por personas de la tercera edad, a partir de un Entorno dotado de Inteligencia Ambiental, acondicionado con sensores, actuadores y una inteligencia proporcionada por una estructura de datos llamada pizarra, donde a través de una detonación (obligatoria) o disparo de un estado de alerta, por ausencia de actividad por ejemplo, habilite el monitoreo de la casa-habitación desde remoto, en busca de la persona vulnerable.

Como limitación, se requiere de la infraestructura de un Entorno con Inteligencia Ambiental. Por ahora se han realizado pruebas en un laboratorio real, dotado de esta tecnología compuesta por sensores, actuadores, y una capa middleware que permite el modelado virtual del mundo.

#### V. PROPUESTA

La propuesta radica en realizar un seguimiento de las actividades del usuario dentro del Entorno Activo, a través de una red de sensores, actuadores y aplicaciones sensibles al contexto. Como premisa, el sistema debe ser lo menos intrusivo posible para evitar que las personas vulnerables se sientan vigiladas, por lo que se pretende evitar el uso de cámaras.

Para que la interacción entre usuario y Entorno Activo sea de lo más natural, se pretende evitar el uso de dispositivos de localización (tokens), proporcionando versatilidad.

En caso de que el Entorno detecte un estado de emergencia, adquirido por la red de sensores o por las aplicaciones sensibles al contexto, por ausencia de actividad o alguna actividad sospechosa, se detonará una alerta, que notificará a un usuario remoto (encargado del cuidado de la persona vulnerable). A partir del estado de emergencia, se habilitará una interfaz que permitirá el uso de un robot móvil desde remoto, que adicionado con una cámara permitirá recorrer el Entorno Activo en busca información de la persona vulnerable.

La propuesta parte de crear un sistema de vigilancia poco intrusivo, a partir del uso de un robot como medio de adquisición del contexto del usuario que sólo podrá ser empleado cuando y sólo cuando ocurra un estado de emergencia desencadenado por el Entorno Activo.

La propuesta pretende primero usar el robot como medio de retroalimentación, de tal manera que una persona vulnerable tendrá conocimiento en todo momento de que la cámara del robot sólo estará activa a partir del estado de emergencia, notificación por demás evidente a partir del movimiento del robot, esto pretende evitar abrumar a un usuario vulnerable con una vigilancia permanente del Entorno Activo, desalentando su resistencia al uso de la tecnologías pervasivas, cuidando su privacidad y estilo de vida independiente. Recíprocamente, se pretende proporcionar tranquilidad a las personas a cargo de la seguridad y bienestar de las personas vulnerables.

#### VI. DESCRIPCION DEL ENTORNO ACTIVO

La propuesta se desarrolla en el ámbito de interfaces de usuario (human Computer Interfaces) aplicadas hacia un área

emergente como lo son los “Entornos con Inteligencia Ambiental”, también llamados “Entornos Activos” (Active Environments), “Entornos Inteligentes” (Intelligent Environments, Smart Environments).

Los Entornos con Inteligencia Ambiental integran muchos y variados dispositivos de cómputo, dispersos e integrados de tal manera que pasan desapercibidos a los usuarios. Estos dispositivos, son capaces de establecer comunicaciones y de ofrecer servicios a través de novedosas interfaces, ayudando al usuario en sus tareas cotidianas de una manera poco intrusiva y sensible al contexto.

AmILabMx (Ambient Intelligence Laboratory-México) se desarrolló a partir del año 2006 con el objetivo de fomentar la investigación en el campo, de los conocimientos hasta entonces desarrollados por la Universidad Autónoma de Madrid, a través del despliegue de un Entorno Activo en el Instituto Tecnológico Superior Zacatecas Norte (ITSZN), México, en el cual la interacción humano-computadora se realice de manera natural y sensible al contexto de la tarea que se está ejecutando.

El Entorno Activo en cuestión, se encuentra actualmente en un estado de desarrollo avanzado y funcional, desenvolviéndose como un laboratorio de pruebas que promueve la colaboración en líneas de investigación afines a la Inteligencia Ambiental y al desarrollo de interfaces de usuario a través de tres escenarios que consisten en un aula de clases, una sala de reuniones y una sala de estar. Hasta ahora, cuenta con dos interfaces que ofrecen control directo sobre el entorno, una de ellas, permite interaccionar con los diferentes electrodomésticos y dispositivos a través del lenguaje natural.

Se ha desarrollado un primer prototipo a través de cinco tecnologías:

- Una arquitectura de hardware basada en tecnología TCP/IP, buses domóticos X10 y dispositivos Phidget que, interconectan los diferentes dispositivos que actualmente se pueden encontrar en entornos domésticos, ofimáticos o industriales, convirtiéndolos de una manera flexible, en integrantes del Entorno Activo.
- Una arquitectura de software de integración basada en agentes autónomos, que a su vez, hacen uso de una arquitectura basada en “pizarra” como mecanismo de coordinación, mismo que realiza el control inteligente de los diversos dispositivos, tanto sensores como actuadores.
- Agentes basados en reglas que proporcionan los medios para hacer que el Entorno se adapte a las necesidades del usuario (control indirecto).
- Un agente de diálogos en lenguaje natural, sensible al contexto del usuario. Este agente se basa en herramientas estándar de reconocimiento y síntesis de la voz, haciendo incidencia en el procesamiento del lenguaje natural en un Entorno multidimensional.



- Una interfaz gráfica a través de Internet que permite controlar el Entorno Activo desde cualquier lugar remoto.

Los dispositivos y aplicaciones desplegados en cada escenario provienen de tecnologías heterogéneas, gobernadas por un capa de contexto que se encarga de abstraer sus diferencias.

#### A. La capa física

La infraestructura física del entorno, está formada por una red de sensores (presencia, temperatura, iluminación, etc.) y actuadores (switches, etc.) comerciales y estandarizados. Una parte de ellos son dispositivos domóticos X10, algunos otros son tecnología tipo Phidget. Hay una red adicional Ethernet, encargada del transporte de datos, audio y video generados por las diferentes fuentes disponibles (cámaras ip, webcams, audio streaming, etc.). La integración de nuevos dispositivos no representa ningún problema, solo basta con conectarlo a una de las dos redes (dependiendo de su naturaleza), y la capa de contexto se encargara de ejercer el control sobre el dispositivo y su información. Por ejemplo, la puerta principal del laboratorio fue equipada con un lector RFID. El acceso al laboratorio se hace pasando una tarjeta por el lector, este envía el código a la capa de contexto donde se realiza la identificación de la persona, procediendo al envío del comando de apertura.

#### B. La capa de contexto

La proliferación de las redes de comunicación y la diversidad de protocolos [4], complican la integración de dispositivos en un entorno. Para el Entorno AmILabMx, se propone una capa de contexto que sirve de interfaz entre los diversos dispositivos de cómputo, hacia la integración de un entorno activo [5]. La capa de contexto empleada, parte de la filosofía de que la representación del mundo se puede hacer a través de un modelo centralizado.

Respecto a su implementación, ésta puede hacerse distribuida, proporcionando la suficiente potencia para representar las interacciones complejas de todo lo que convive en el entorno. Además, la capa de contexto proporciona un modelo unificado del mundo, abstrayendo los detalles de comunicación y las diferencias que existen entre los dispositivos del entorno.

La implementación del modelo resulta en una estructura de datos global llamada "pizarra" [6]. La pizarra es un repositorio donde se almacena una representación de todo lo que existe en el entorno, incluyendo a los usuarios. Por lo tanto, cada dispositivo, usuario y recurso del entorno activo estará representado mediante una entidad en la pizarra. Además, contiene una representación del flujo de información existente entre los dispositivos físicos (micrófonos, altavoces, cámaras, pantallas, etc.).

La información de la pizarra es utilizada por los diferentes dispositivos para comprender el contexto y adaptarse a éste. Cada pizarra es un servidor que puede ser accedido mediante el protocolo cliente-servidor TCP/IP. HTTP ha sido escogido como el protocolo de transporte por su simpleza y amplia cobertura. Para el intercambio de información entre las

aplicaciones y la pizarra en el servidor, se emplea el lenguaje XML.

La representación del contexto contempla algunas de las ideas expresadas por Dey en [7], para representar personas, lugares y recursos. Esto se hace mediante la definición de conceptos (o clases) que sirven como plantillas, a partir de las cuales son creadas instancias, también llamadas entidades. Cada entidad pertenece a un concepto, y se representa mediante un nombre y un conjunto de propiedades. Cada propiedad tiene un nombre y un valor que puede ser una literal o alguna otra entidad. En este caso, se establece una relación entre la primer entidad y la segunda. Las relaciones son unidireccionales, aunque para cada relación siempre se puede definir la inversa. Finalmente, se pueden definir conjuntos de parámetros que pueden ir asociados tanto a un concepto, a una entidad, a una propiedad o a una relación.

La capa de contexto implementada en la estructura de datos llamada pizarra, proporciona el soporte requerido por las aplicaciones. Se han implementado dos interfaces de usuario en el entorno activo: una interfaz gráfica que funciona a través de Internet [8] y que permite controlar los dispositivos de la habitación, la segunda, es un agente de diálogo en lenguaje natural [9] que permite que el usuario interactúe con el entorno. Ambas, son configuradas dinámicamente a través de la información proporcionada por la pizarra.

#### C. Agentes basados en reglas

Una vez que se ha desarrollado una estructura de datos tipo pizarra, capaz de representar al mundo y de dotar al entorno de una visión global, aglutinando y homogeneizando la información aportada por los distintos dispositivos del entorno, se plantea otro reto, que es cómo transformar un Entorno Perceptivo (percibe el mundo que lo rodea) en Entornos Interactivos, en los que sus habitantes pueden comunicarse e interactuar con él, capaces de tomar decisiones basadas en la información de contexto.

Para lograr este comportamiento interactivo, se ha implementado un sistema de agentes cuyo fin es dotar de un mecanismo mediante el cual los diferentes usuarios puedan comunicar sus preferencias sobre cómo abordar necesidades de comportamiento automático, a fin de modelar el entorno a su gusto.

La arquitectura de este componente de software se basa en un conjunto de módulos independientes entre sí (los agentes), cada uno de los cuales contiene un conjunto de reglas que modelan las necesidades específicas de cada usuario sobre el comportamiento del entorno. Cada regla está constituida por tres elementos [10]: El detonante que define la propiedad de una entidad supervisada y dispuesta en el entorno, responsable de activar la regla. Las condiciones, que especifican las condiciones del contexto sobre las cuales la regla es considerada válida para su ejecución. Y por último la acción, que describe los cambios que deben suceder en el entorno dadas las condiciones del contexto. Así pues, una regla puede expresar que cuando la puerta se abra (el detonante), si la luz está apagada (la condición), se encienda la luz (la acción).

Las reglas se especifican mediante un lenguaje diseñado para garantizar la expresividad, explicación y aprendizaje del

modelo. El lenguaje puede ser fácilmente leído por el usuario y permite describir el contexto y las acciones que se deben realizar en él. De esta manera, por ejemplo, la regla expresada en el párrafo anterior, podría ser definida en éste lenguaje de la siguiente manera:

```
#Regla: cuando se abre la puerta principal, si la luz está
apagada entonces que se encienda
device:main_door:status::device:lamp_1:status=OFF;
device:main_door:status=1->device:lamp_1:status:=ON
```

Uno de los requerimientos más importantes que trata de cubrir el modelo, es el de no intrusión. Los agentes tienen mecanismos de configuración que permiten la aplicación de las reglas en un modo supervisado por el usuario (preguntando por la aprobación del usuario para la ejecución de una regla) o bien en un modo autónomo (aplicando las reglas sin esta aprobación). Además, cada agente ejecutado en el sistema es considerado una entidad en la pizarra de contexto, los cuales pueden ser activados o desactivados en cualquier momento dependiendo de las necesidades de los usuarios del entorno.

AmILabMx, se ha dotado de un mecanismo mediante el cual los diferentes usuarios puedan comunicar sus preferencias sobre cómo abordar cuestiones cotidianas (por ejemplo: cuando encienda la Tv, apaga la luz), diseñando ellos mismos sus propias aplicaciones. Este mecanismo está enfocado principalmente a proporcionar soluciones al problema de la adaptación, dotando al usuario de la capacidad de modelar el entorno a su gusto.

La manera de construir agentes, es a través de un lenguaje orientado a la creación de reglas. Cada regla queda dividida en tres partes: detonante, condiciones y acción. El lenguaje para la creación de reglas, fue construido bajo la premisa de sencillez y adecuación.

La primera, en cuanto a facilitar la tarea de trasladar el sistema al terreno de interfaces más complejas como por ejemplo: la interacción mediante lenguaje natural. La segunda, que sea capaz de expresar los deseos humanos sobre el entorno.

Así pues, un agente está constituido por un conjunto de reglas. En base a ellas, y a los cambios producidos en el contexto, reacciona produciendo una salida en forma de órdenes o cambios a realizar sobre ese mismo contexto. La forma en que reacciona a esos cambios, las consecuencias internas que conllevan así como el proceso de inferencia que detonan puede representarse en forma de algoritmo.

La versatilidad de las reglas es tal, que permite la generación de los estados de emergencia propuestos en la sección V.

## VII. INTERFACES DE USUARIO

El Entorno Activo cuenta con una interfaz de diálogos en lenguaje natural [9], cuya importancia radica en la manera de apoyarse en la información contextual almacenada por la pizarra, haciendo posible establecer una conversación coherente relacionada con el control de los dispositivos y la información de los usuarios. La interfaz ejecuta varios diálogos relativos al Entorno que compiten por ser el más idóneo para la conversación actual. Un supervisor de diálogos, tiene la tarea

de elegir el diálogo más adecuado de acuerdo a la entrada de datos proporcionada por el usuario (proporcionada por el reconocedor de voz) y la información contextual proveniente de la pizarra. Además, el supervisor de diálogos tiene la tarea adicional de activar y desactivar diálogos en función de si tienen o no sentido.

Cada diálogo, se centra en una tarea específica, por ejemplo, el diálogo de las luces se encarga de controlar el estado encendido o apagado de las lámparas de una habitación. Esto es, el diálogo tiene la capacidad de leer y escribir sobre los valores de las lámparas almacenadas en la pizarra.

Debido a que la fase de reconocimiento de voz no es del todo exacta, la información de contexto proveniente de la pizarra toma un rol muy importante en las decisiones del supervisor, esto es, cada sentencia que proviene del usuario, puede venir con ambigüedades, y el supervisor las resuelve empleando la información de la pizarra. Otra opción, es ofrecer soluciones al usuario dependientes del contexto. Por ejemplo, si el reconocedor de voz solo ha detectado con precisión la palabra luz, el supervisor podrá corroborar el estado de las luces. Si estas están apagadas, directamente puede ofrecer la opción de encenderlas.

Además del agente de diálogo en lenguaje natural, el laboratorio cuenta con una interfaz gráfica que es desplegada por cualquier navegador a través de un applet y desarrollada para controlar todos los dispositivos y electrodomésticos del Entorno Activo. Esta aplicación se puede considerar como una vista parcial de la información almacenada en la estructura de datos llamada pizarra. Al iniciar la aplicación, hace una lista de las habitaciones del laboratorio (cuenta como habitación cada escenario en el que está dividido el laboratorio), y para cada habitación, genera un mapa que incluye una representación gráfica de cada dispositivo que se encuentra físicamente en el Entorno, así como su localización. Cada vez que el usuario hace clic con el ratón sobre la imagen de un dispositivo, aparece una interfaz (panel de control) que controla el funcionamiento de ese dispositivo (por ejemplo: si hace clic sobre una lámpara, aparece un botón que la enciende o la apaga).

La pizarra, contiene información genérica respecto al número de habitaciones, así como de los dispositivos con los que cuenta cada una de ellas. La representación de cada dispositivo incluye las propiedades requeridas para su control, imagen a desplegar en el mapa, coordenadas donde se debe posicionar su imagen, además de otras posibilidades útiles a otras aplicaciones. De esta manera, cada vez que se inicia la interfaz gráfica, se hace un recuento de los dispositivos que pertenecen a la habitación que se quiere controlar, y se genera dinámicamente su respectivo mapa.

El panel de control para cada dispositivo se genera dinámicamente a partir de sus propiedades. Además, se han definido un conjunto de widgets genéricos que consisten de cajas de texto, botones, barras de desplazamiento, entre otros, de tal manera que cuando se quiere controlar algún dispositivo a través del applet, se leen las propiedades de su descripción existente en la pizarra, y las propiedades son traducidas a widgets generando su correspondiente panel de control.

Los agentes de diálogos y la interfaz gráfica son aplicaciones completamente compatibles, esto es, pueden coexistir y funcionar simultáneamente. Ambas aplicaciones muestran dos maneras diferentes de controlar el Entorno. De igual manera, se pueden seguir añadiendo nuevas interfaces, sin tener que realizar modificaciones en la información de contexto almacenada en la pizarra.

### VIII. SITUACIÓN ACTUAL

Actualmente se cuenta con un laboratorio que reproduce la visión de Mark Weiser [11] a través de varios escenarios. Sus aplicaciones sensibles al contexto proponen varias interfaces multimodales para realizar la interacción con el Entorno de Inteligencia Ambiental. Algunas de ellas emplean un control directo sobre los dispositivos del entorno (interfaz gráfica, el gestor de diálogos, entre otros), y también se cuenta con una interfaz de control indirecto, misma que a través de reglas permite dotar de un comportamiento reactivo al Entorno Activo. Esta interfaz es la que actualmente se emplea para detonar reglas del tipo de alertas (que bien pueden ser un correo electrónico, SMS, etc.).

También, se creó un robot sensible al contexto, esto es, las mismas aplicaciones que permiten ejercer un control directo sobre el Entorno Activo, permiten el control del robot, haciendo posible la supervisión remota de los usuarios en caso de alguna emergencia médica.

Como trabajo restante, falta crear algún mecanismo que permita visualizar de manera simple y poco intrusiva la actividad del entorno, tal como lo hace The Aware Home [3] con su portaretrato digital. Los resultados en el laboratorio son alentadores con respecto a la emisión de alertas y al uso del robot como medio de supervisión de las personas vulnerables una vez que se quedan solas.

### IX. CONCLUSIÓN Y TRABAJO FUTURO

Este sistema permite a los usuarios vulnerables, interactuar con un Entorno Activo en presencia de necesidades particulares y básicas para el cuidado de su salud. También realiza un monitoreo continuo del estado del Entorno, para la detección de posibles cambios no habituales. Esto se logra a través de simples reglas que interactúan con una gama de modalidades, dando mayor nivel y prioridad a la asistencia sanitaria, y demandas de rutinas y necesidades habituales.

Una desventaja es la resistencia al cambio que suponen estas tecnologías para personas de la tercera edad. Es necesario recordar que los usuarios del sistema serán personas mayores posiblemente sin experiencia en tecnologías de la información, de diferentes orígenes sociales y culturales, y con distinto nivel educativo.

La computación ubicua y los Entornos dotados de Inteligencia Ambiental hacen posible poner la tecnología al servicio de los más vulnerables, y de quienes tienen a su cargo su supervisión, por lo que a futuro será necesario trabajar en algún dispositivo que permita la recepción de las alertas de

seguridad del Entorno Activo de una manera poco intrusiva (imagínese una pulsera sincronizada a su teléfono móvil en donde de acuerdo al color de un led, le notifique acerca de un estado de alerta del Entorno Activo remoto).

El sistema descrito es actualmente una línea de investigación implementada dentro del laboratorio AmILab (Laboratorio de Inteligencia Ambiental).

### AGRADECIMIENTO

Este proyecto de investigación es posible gracias al patrocinio del Instituto Tecnológico Superior Zacatecas Norte y Tecnológico Nacional de México, a través de los proyectos 358.15-PD y 359.15-PD. Además, nuestro reconocimiento al apoyo del Laboratorio de Inteligencia Ambiental de la Universidad Autónoma de Madrid.

### REFERENCIAS

- [1] A. R. Delgado, A. Robinet, J. McGinn, V. Grout and R. Picking, "Assistive human-machine interfaces for smart homes," in SEIN 2007: Proceedings of the Third Collaborative Research Symposium on Security, E-Learning, Internet and Networking, 2007, pp.145.
- [2] G. Demiris, M. J. Rantz, M. A. Aud, K. D. Marek, H. W. Tyrer, M. Skubic and A. A. Hussam, "Older adults' attitudes towards and perceptions of smart home technologies: a pilot study," Medical informatics and the Internet in Medicine, vol. 29, num. 2, pp. 87-94, 2004.
- [3] C. D. Kidd, R. Orr, G. Abowd, C. G. Christopher, I. A. Essa, B. MacIntyre, E. D. Mynatt, T. Starner and W. Newstetter, "The Aware Home: A Living Laboratory for Ubiquitous Computing research," in Proceedings of the Second International Workshop on Cooperative Buildings, Integrating Information, Organization, and Architecture, pp. 191-198, 1999.
- [4] P. A. Haya, X. Alamán and G. Montoro, "A comparative study of communication infrastructures for the Implementation of Ubiquitous Computing," UPGRADE, The European Journal for the Informatics Professional, 2, 5, pp. 26-40, 2001.
- [5] P. A. Haya, G. Montoro and X. Alamán, "A Prototype of a Context-Based Architecture for Intelligent Home Environments," On the Move to Meaningful Internet Systems 2004: CoopIS, DOA, and ODBASE, OTM Confederated International Conferences, Proceedings, Part I, pp. 25-29, 2004.
- [6] R. S. Englemore and A. J. Morgan, "Blackboard Systems," Addison-Wesley, 1988.
- [7] A. K. Dey, "Understanding and Using Context," Personal Ubiquitous computing, Vol. 5 (1), pp. 4-7, 2001.
- [8] X. Alamán, R. Cabello, F. Gómez-Arriba, P. A. Haya, A. Martínez, J. Martínez and G. Montoro, "Using context information to generate dynamic user interfaces," Proceedings of the Tenth International Conference on Human-Computer Interaction, Vol. 2, pp. 345-349, 2003.
- [9] G. Montoro, X. Alamán and P. A. Haya, "Spoken Interaction in Intelligent Environments: A Working System," Advances in Pervasive Computing, 2004.
- [10] M. García-Herranz, P. A. Haya, A. Esquivel and X. Alamán, "Easing the Smart Home: Semi-automatic Adaptation in Perceptive Environments," Journal of Universal Computer Science, Vol. 13 (12), pp. 1-10, 2007.
- [11] M. Weiser, "The Computer for the 21<sup>st</sup> Century," SIGMOBILE Mob. Comput. Commun. Rev. Vol. 3 (3), pp. 3-11, ACM 1999.

# Prototipo de sistema de telemetría de un satélite CanSat

Ing. Feliciano Hernández Luis Fernando  
ESCOM-IPN  
Ciudad de México  
[lfer091@gmail.com](mailto:lfer091@gmail.com)

Ing. Mariano Arumir Rivas  
Academia de Ingeniería  
UAEH  
Pachuca, Hidalgo  
[mar317@hotmail.com](mailto:mar317@hotmail.com)

M. en C. Jaime Hugo Puebla Lomas  
Ingeniería en Sistemas Computacionales  
ESCOM-IPN  
Ciudad de México  
[jpuebla@ipn.mx](mailto:jpuebla@ipn.mx)

**Abstract**—In this article we will focus on showing the results obtained with the use of a CANSAT the testing for measuring temperature, atmospheric pressure, altitude and distance from the device to the ground. Also made calculations related to the telemetry to a distance of 800km, but in this case we development the calculations considering the tool of the extrapolation. All the calculations will be considering into Telemetry, Calculations of Satellite Link. The study of the case we considering characteristic of the nanosatellite Cube Sat

**Keywords**—Telemetry, Satellite, Instrumentation, Software

## I. INTRODUCCIÓN.

En el año 1999 la Universidad Estatal Politécnica de California, desarrolló e implemento un nano satélite al que llamo CubeSat con la complejidades físicas, electrónicas y mecánicas que representa un satélite real, dicho satélite tiene un volumen exacto de 1 litro y con un peso menor a 1.33 Kg. Se encuentran ubicados en órbita baja es decir no más de 800 Km de altura. El CubeSat cuenta con diferentes subsistemas como son: La carga útil, comandos y datos, la fuente eléctrica, control térmico, posicionamiento y propulsión, comunicaciones y la computadora de vuelo <sup>[1]</sup>.

El desarrollo espacial en México ha sido lento, pero hace 10 años el Dr. Esaú Vicente Vivas, profesor e investigador de la UNAM especializado en el desarrollo e implementación de Sistemas Espaciales con el apoyo de otras universidades e instituciones importantes del país, desarrollaron un prototipo de un microsátélite al que denominaron “SATEDU” <sup>[2]</sup>.

En el presente trabajo se realizaron pruebas de comunicación entre un Equipo Terminal de Datos (PC) y un satélite CanSat, empleando la extrapolación como herramienta para garantizar el enlace de comunicación tomando en cuenta la sensibilidad de la tarjeta de comunicación, así como el uso de la ecuación de Friis.

La extrapolación se realizo tomando en cuenta la distancia del enlace de comunicación entre una estación terrena y un nano satélite de órbita baja, de forma real esto debido a que no se cuenta con la infraestructura de comunicación satelital. Por lo que con el supuesto de tenerlos y con ciertas características se realizaron los cálculos de telemetría para una distancia de 800km. Para ello se debieron hacer los cálculos correspondientes para el enlace satelital tanto de subida como el del enlace de bajada <sup>[3]</sup>; tomando en cuenta variables físicas como son la temperatura, las condiciones atmosféricas, dimensiones de la antena para así determinar la frecuencia de operación y la potencia de transmisión. Y de forma práctica con la técnica de extrapolación como se menciono, se realizó la comunicación a 7 metros, con la utilización de 2 Xbee S1.

Aunque la estructura física y eléctrica, así como el tamaño, costo y funcionalidad por el cual fueron hechos de un CanSat con respecto a un CubeSat son muy diferentes, el CanSat permite obtener experiencia acerca de la tecnología y ciencia espacial, a una escala muy pequeña incomparable con la de un CubeSat pero útil al momento de realizar pruebas de medición, comunicación, estabilidad, etc.

Para llevar a cabo la adquisición de datos será necesario desarrollar una interfaz entre las variables a medir y los sensores. Para ello se ocuparon el software de Lab-VIEW para la visualización de la información en forma gráfica y el IDE de Arduino para la programación del microcontrolador. Ver Figura 1.

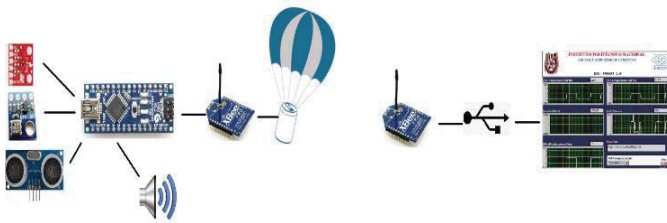


Figura 1. Esquema general del Sistema Híbrido

## II. METODOLOGÍA

En el presente apartado se explicarán las partes importantes para el desarrollo de este prototipo.

Comenzamos con la telemetría a 800km; para ello se requieren la ecuación (1) [4] con los datos supuestos mostrados en la Tabla 1:

$$(CN)_{up} = PIRE_{sat} (dBW) - 20 \log(4\pi f d u c) + G_u T_u (dBK) - 10 \log(k) - 10 \log(B) - BOI (dB) - L (dB) \dots \dots \dots (1)$$

Tabla 1. Parámetros requeridos en la Telemetría.

Parámetros de enlace	Ascendente	Descendente
<b>PIRE del Transmisor</b>	21.5 dBW	6 dBW
<b>Distancia mínima</b>	800 km	800 km
<b>Distancia máxima</b>	1362.7 Km	1362.7 Km
<b>Frecuencia Portadora (rango de la tarjeta es de 433.05 – 434.79)</b>	433.70 MHz	434.40 MHz
<b>Ganancia de antena Trasmisora</b>	11.5 dBi	3 dBi
<b>Perdidas por distancia Max.</b>	127.5 dB	136.2 dB
<b>Perdidas por distancia Min.</b>	122.7 dB	131.3 dB
<b>Perdidas varias</b>	3 dB	3 dB
<b>Ganancia de Antena Receptora</b>	2 dBi	14 dBi
<b>Constante</b>	Nombre	Valor
<b>c</b>	Velocidad de la luz	299,792,458 m/s
<b>T</b>	Temperatura ambiente	298 °K
<b>k</b>	Constante de Boltzmann	1.3805488 x 10 <sup>-23</sup> J/K

Los resultados para enlace satelital de subida así como de bajada para una distancia máxima de 1200km y a una distancia mínima de 800km, se muestra a continuación:

$$(CN)_{up} 800km = -32.0958 \text{ dB}$$

$$(CN)_{up} 1200km = -41.5225 \text{ dB}$$

$$(CN)_{down} 800km = -53.7099 \text{ dB}$$

$$(CN)_{down} 1200km = -63.2365 \text{ dB}$$

Esos dBs son la cantidad de potencia que requiere la estación terrena para poder realizar la comunicación a esa distancia. Para ganar más potencia es necesario aumentar la ganancia de las antenas, adquirir un amplificador de potencia o seleccionar una tarjeta de comunicación potente con una buena ganancia.

Por cuestiones de presupuesto se opto por utilizar los dispositivos de comunicación Xbee Pro, los cuales operan en una frecuencia de operación de 2.4GHz, usando la tecnología de comunicación Zig Bee, en lugar de las tarjetas de Radiometrix TR21-433-5

Posteriormente se estudió el concepto de la extrapolación que consiste en realizar un análisis a partir de una escala pequeña de un sistema de comunicación para poder extrapolarlo a una distancia considerable; para ello se utilizó la fórmula de Friis [5] y con las características técnicas del Xbee- S1.

$$LS = 20 \log[4\pi f d c] (dB) \dots \dots (2)$$

Se tiene que  $f = 433 * 106 \text{ Hz}$ ,  $d = 6 \text{ m}$ . Sustituyendo en (2) se obtiene que:

$$LS = 20 \log[4\pi(433 * 106 \text{ Hz})(6 \text{ m})299792548 \text{ m/s}] = \mathbf{40.74 \text{ dB}}$$

Con ello se obtiene la cantidad de potencia que requiere este sistema para comunicarse a 6 metros y a partir de ahí extrapolarlo a la distancia de 800 km.

Posteriormente continuamos con el desarrollo electrónico, utilizando los siguientes sensores:

### a) Sensor de temperatura TMP102:

Es un sensor de temperatura digital que se usó en el CubeSat llamado ARDUSAT [6] en el cual cumplió con su misión en 2012. Dicho sensor ofrece una precisión 0.0625°C. Mide 1.6mm x 1.6mm, es compatible con la interfaz I2C, y requiere un voltaje entre los 1.4V a 3.6 V con un uso de corriente de 10uA a 1uA. Funciona a temperatura ambiente entre -40°C a 125°C, Ver Figura.2

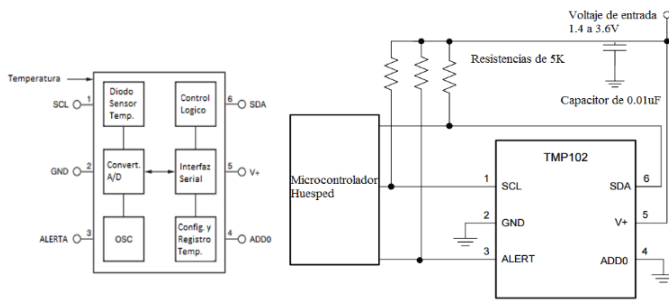


Figura 2. Configuración electrónica del sensor TMP102.

b) Sensor de presión atmosférica y temperatura BMP180:

El dispositivo BMP180 [7] es un sensor digital de presión, tiene una sensibilidad de 0.01hPa y 0.1°C para la temperatura por cada 4ms de lectura, Son operables bajo una temperatura de -40°C a 65°C.

Requiere una alimentación de voltaje de 1.8V o 1.62V a 3.6V y con mínimo de corriente de 5uA hasta 1A. Sus medidas son 3.6mm x 3.8mm. Para su configuración electrónica ver la Figura3.

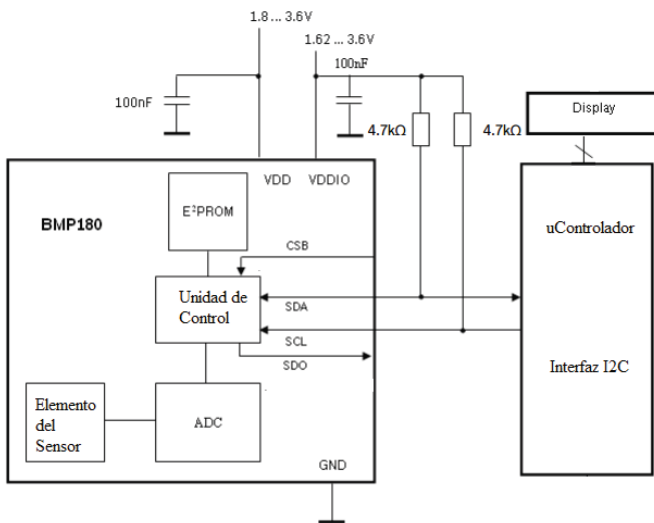


Figura 3. Diagrama electrónico para la conexión del sensor.

c) Sensor Ultrasonico HC-SR04 :

El módulo HC-SR04 [8] provee un rango de medición entre los 2cm a 400 cm [no-contacto], el rango de presión puede alcanzar los 0.3cm. Requiere 5V de alimentación, 15mA de corriente, una frecuencia de 40Hz y un pulso de entrada [Trigger] y lanza un pulso [Echo] de salida. Solo tiene un ángulo de medición de 30 grados. Para que comience a medir,

el [Trigger] del SRO4 tiene que recibir un pulso alto de 5V por los últimos 10us, esto inicializa el sensor para poder transmitir una señal ultrasónica cada 8 ciclos a 40kHz, y esta es reflejada al momento de impacto.

Una vez que se han explicado los sensores más importantes para la medición de las variables físicas del CanSat se proseguirá con el diseño electrónico. En la Figura 4 se muestra el diagrama esquemático del CanSat, el cual se compone de un nano Arduino versión 2.0, el sensor ultrasónico HC-SR04, el sensor TMP102, así como el sensor BMP180, una fuente regulada que va integrada una pila de 9v para así regularlo a 5v, las salidas de la bocina y de un interruptor, la configuración particular del Xbee para recibir datos y por último dos leds para indicar la distancia del CanSat a la tierra.

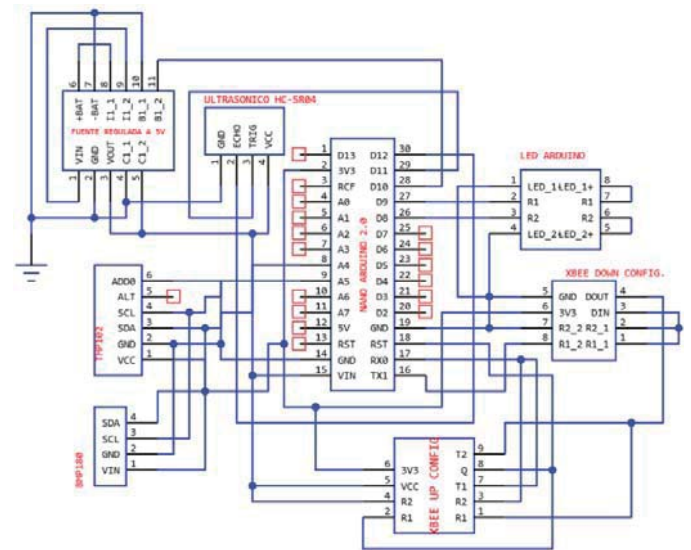


Figura 4. Diagrama Esquemático del CanSat.

Cabe recalcar que la programación del nano Arduino se realizó con el IDE de Arduino y con las librerías correspondientes a los sensores que se mencionan en el presente trabajo.

Una vez terminado el circuito electrónico, se prosiguió con la elaboración del software. Para ello se usó Lab-VIEW versión 2014 y se realizó la siguiente programación, Ver Figura 5.

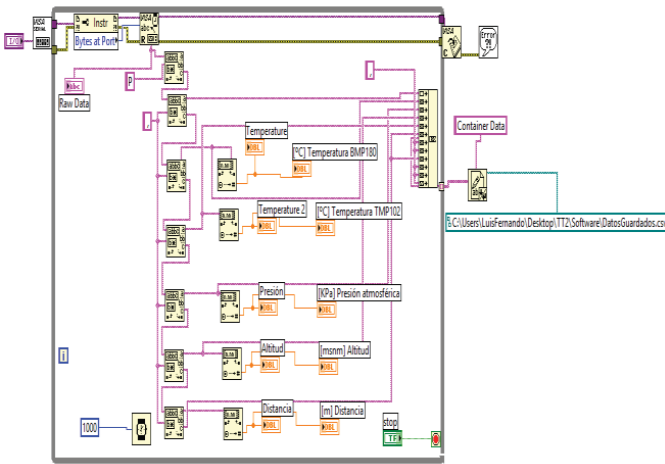


Figura 5. Diagrama General del Sistema de adquisición de datos, desarrollado en LabVIEW.

Se separa la trama enviada por el microcontrador del nano Arduino, para posteriormente mostrarlo en los diferentes campos de medición y concatenarlo en un archivo .txt.

### III. RESULTADOS

El resultado de la programación con Lab-VIEW se obtuvo la siguiente interfaz gráfica donde se pudieron visualizar los datos con gráficas y números de punto flotante. Ver Figura 6.



Figura 6. Interfaz gráfica previa a su funcionamiento.

La construcción final del CanSat quedó como se muestra en la Figura 7 para así realizar mediciones en diferentes partes y visualizarlo en la interfaz, ver Figura 8. dicha información se muestra en la Tabla 2 por los días que se realizaron las diferentes pruebas.

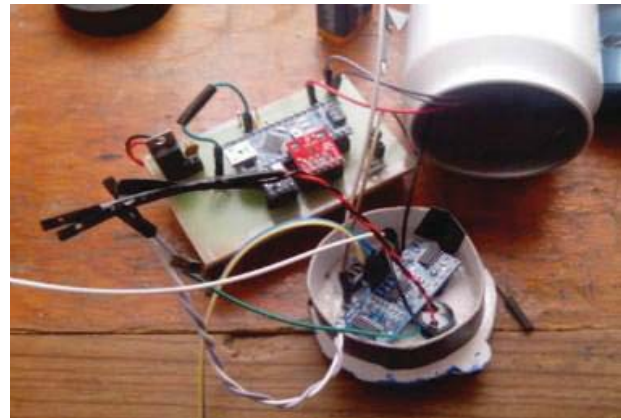


Figura 7. Sistema de Comunicación y Telemetría del CanSat.

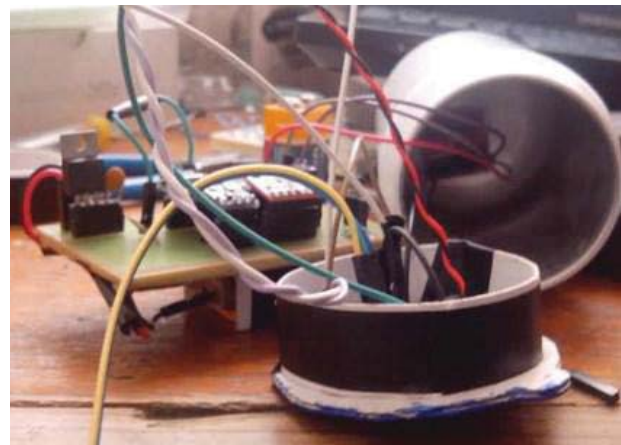


Figura 8. Sensores de proximidad del CanSat.

En la Figura 9, se muestra la operación de la interfaz con cada uno de las variables con las que se realiza la telemetría, empleando el lenguaje de programación de LabVIEW.



Figura 9. Interfaz para la Telemetría del CanSat.

En la Tabla 2 se muestran los diferentes escenarios donde se realizaron cada una de las mediciones de telemetría.

Tabla 2. Resultados obtenidos en diversos escenarios

Fecha	Lugar	Ta.1 °C	Ta.2 °C	Presión atmosférica	Altitud msnm	Distancia (m)	Tiempo Encendido de
07/08/2015	Casa	22.3	22.1	78	2140	10	4 min.
14/08/2015	Casa	21.2	21.1	77	2140	5	3 min.
21/08/2015	Sociedad Astronómica de México	23.7	23.4	73	2095	20	6 min.
28/08/2015	Sociedad Astronómica de México	22.6	22.4	75	2095	10	6 min.
04/09/2015	UAEH	20.6	20.4		2180	4	10 min.
11/09/2015	ESCOM	21.5	21.7	76	2099.67	4	4 min.
18/09/2015	ESCOM	22.8	23.1	77.5	2100	10	5 min.

#### IV. CONCLUSIONES

La experiencia adquirida en esta investigación se ve reflejada en el correcto funcionamiento del CanSat para la medición de variables tales como; temperatura, presión atmosférica, la altitud, y la distancia del aparato con respecto al suelo. Teniendo en cuenta problemas a resolver como el espacio, la velocidad con la cual se envía o se reciben datos, la construcción de un circuito electrónico capaz de ser introducido dentro de una lata de refresco, enviar toda esta información con un Xbee, aprender diferentes tecnologías para desarrollar cada uno de los componentes del presente trabajo. Aunque aún está en un fase de desarrollo para ser parte de un satélite real, es un buen punto de partida para quienes se encuentren interesados en la ciencia y tecnología espacial.

#### RECONOCIMIENTOS

Los Autores agradecen a la Escuela Superior de Ingeniería Mecánica y Eléctrica, Escuela Superior de Cómputo del Instituto Politécnico Nacional, a la Sociedad Astronómica de México así como la UAEH por el apoyo recibido y las facilidades otorgadas para el desarrollo del presente Proyecto.

#### REFERENCIAS:

- [1] Wiley J. Larson, James R. Wertz, Space Mission Analysis and Design, Third Edition, 2005.
- [2] Esaú Vicente-Vivas, E.A. Jiménez, Z.L. Carizales, R. C.A. Sánchez, Córdova, R. Alba, M. A. García and G. Islas, "Successful Development of a Portable Didactic Satellite for Training and Research in Satellite Technology", CORE-2009, 10th Computing Congress, CIC-IPN, México City, May 27-29, 2009.
- [3] Ing. Mariano Arumir Rivas, "Calculo de enlace Satelital" Perito número 316 en Telecomunicaciones especialidad en Comunicación Espacial y Radiocomunicación.

[4] E. Vicente-Vivas, F. Garcia- notetti y Mendieta Jimenez F. J. "Automatic Maintenance payload on board Mexican LEO Microsatellite"; Revista Acta Astronautica Journal, Elsevier Science.

[5] Carlos Rosado "Comunicación por satélite, principios, tecnologías y sistemas" AHCIET, 2000. I.S.B.N.:84-87-644-42-2.

[6] Vicente Vivas Esaú, "Red de Área Local tolerante a Fallas Aplicada a Sistemas de Tiempo Real con Altos Requisitos de Confiabilidad". Revista Científica de la ESIME, Núm. 15, pp. 13-24, Mayo Junio de 2005

[7] Esau Vicente Vivas and Fabian Garcia Nocetti "Reconfigurable Semi- Virtual Computer Architecture for Long Available Small Space Vehicles" Proceedings 2001 IEE Aerospace Conference, Big Sky, Montana USA, March 2001

[8] José Rodrigo Córdova Alarcón, Estimación y Control de Orientación para el Nanosatélite HumSat México, Tesis de Maestría, Universidad Nacional Autónoma de México, 2011.



# Sincronización del bursting de la célula beta mediante redes complejas

D.K. Guevara-Flores\*, J.M. Muñoz-Pacheco\*, E. Zambrano-Serrano†, and O. Félix-Beltrán\*

\*Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, Puebla, México, C.P. 72570

†Departamento de Matemáticas Aplicadas, IPICYT, San Luis Potosí, México, C.P. 78216

Email: jesusm.pacheco@correo.buap.mx

**Resumen**—Las células beta pancreáticas se encuentran agrupadas en los islotes de Langerhans y son las encargadas de producir y segregar insulina cuando su potencial de membrana se sincroniza en una dinámica no lineal denominada *bursting*. Por lo tanto en este trabajo se diseña una red compleja del tipo mundo pequeño para sincronizar células beta aisladas y en consecuencia, analizar la generación de *bursting*. En específico, cada nodo de la red compleja se define mediante el modelo de Pernarowski para una célula beta. Simulaciones numéricas demuestran que la red compleja propuesta es capaz de lograr la sincronización entre diversos tipos de *bursting*.

## I. INTRODUCCIÓN

Las células beta se encuentran en el páncreas y son las encargadas de generar y segregar insulina en el torrente sanguíneo cuando aumentan los niveles de glucosa para que esta pueda ser aprovechada por las demás células del cuerpo [1–4]. Las células beta no trabajan independientemente, sino que se encuentran agrupadas por miles en los denominados islotes de Langerhans. Dentro de cada islote, las células beta sincronizan su potencial membrana en oscilaciones características que se denominan *bursting cuadrado*. El *bursting cuadrado* es un comportamiento periódico que se caracteriza por oscilaciones complejas rápidas seguidas de un periodo de inactividad denominado fase silenciosa. Existen diversos modelos matemáticos que describen la dinámica de la célula beta, entre ellos el propuesto por Pernarowski [1]. Este modelo permite, por medio de cambios en los parámetros de las ecuaciones diferenciales acopladas que lo componen, obtener cuatro tipos distintos de *bursting* además del *bursting cuadrado*, característico de la célula beta. El modelo de Pernarowski permite emular la dinámica de una célula beta aislada, por lo que es posible emplearlo para caracterizar los nodos de una red compleja y analizar el funcionamiento de una parte del islote de Langerhans, así como el comportamiento colectivo emergente de sistemas no lineales aislados al agruparse.

Existen muchas topologías de redes complejas, entre las cuales la llamada *mundo pequeño* es la que más se asemeja a los sistemas biológicos reales, ya que no es ni completamente aleatoria ni completamente regular y no se encuentran nodos aislados del resto, además de presentar un camino corto característico [5, 6].

En este trabajo se diseña una red compleja de tipo mundo pequeño de cinco nodos para estudiar la sincronización de distintos tipos de *bursting*, tales como *bursting cuadrado*, *beating*, *parabólico*, *semiparabólico* y *tapered*. Lo anterior nos permite

el análisis de una dinámica nueva que toma características intermedias entre las señales originales previas a la señal de control. También se explora la topología modificando la posición de los sistemas caracterizados dentro de la red, lo que permite evaluar la importancia del grado de nodo, que intuitivamente su presupone más relevante dado que es el que cuenta con un mayor número de conexiones.

## II. MODELO DE LA CÉLULA BETA

Las células beta presentan un patrón complejo de oscilaciones en el potencial de la membrana denominado *bursting cuadrado* y existe un gran número de modelos matemáticos que lo describen, entre los cuales el descrito por Pernarowski es de tipo fenomenológico, es decir, refleja características cualitativas de las observaciones experimentales realizadas y es lo más simple posible [4]. Este modelo permite simular el comportamiento de una sola célula beta dentro del islote de Langerhans, es adimensional y no lineal. El modelo está compuesto por las tres ecuaciones diferenciales acopladas siguientes:

$$\begin{aligned}\dot{u} &= f(u) - w - c, \\ \dot{w} &= w_\infty(u) - w, \\ \dot{c} &= \epsilon(h(u) - c),\end{aligned}\tag{1}$$

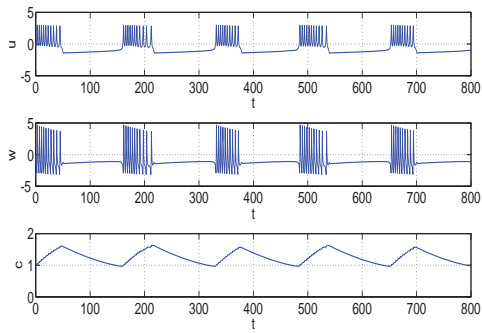
donde  $u$  es el potencial de membrana,  $w$  es el voltaje del canal de activación de potasio y  $c$  la concentración de sustancias que regulan el *bursting*. Las funciones  $f(u)$ ,  $w_\infty(u)$  y  $h(u)$  están descritas por:

$$\begin{aligned}f(u) &= -\frac{a}{3}u^3 + a\hat{u}u^2 + (1 - a(\hat{u}^2 - \eta^2))u, \\ w_\infty(u) &= \left(1 - \frac{a}{3}\right)u^3 + a\hat{u}u^2 - (2 + a(\hat{u}^2 - \eta^2))u - 3, \\ h(u) &= \beta(u - u_\beta),\end{aligned}\tag{2}$$

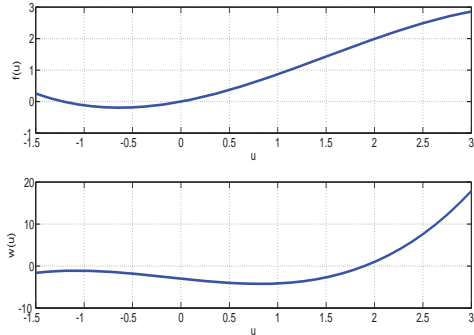
donde los parámetros pueden ser ajustados para obtener comportamientos muy variados. La no linealidad del sistema proviene de los dos polinomios  $f(u)$  y  $w_\infty(u)$  que son de tercer grado.

En la Figura 1(a) se muestra la dinámica de las tres variables de estado, pero la variable de estado  $u$  es la que corresponde a las oscilaciones en el potencial de membrana de la célula y es en la que está enfocado el análisis reportado en este trabajo.

Una de las características más importantes del modelo matemático de la célula beta propuesto por Pernarowski es



(a)



(b)

Figura 1. Dinámica de (a) las variables de estado y (b) los polinomios de tercer grado dentro del rango dinámico de la variable  $u$ .

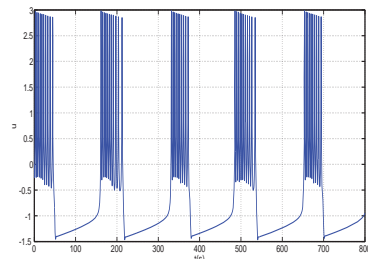
la dinámica tan diversa que presenta ante cambios en los valores de sus parámetros a pesar de la relativa simplicidad de las ecuaciones. Cinco distintos tipos de bursting pueden ser representados con este modelo, el bursting cuadrado característico de la célula beta y otros cuatro, denominados beating, parabólico, semiparabólico y tapered [2].

En la Figura 2(a) se muestra el bursting cuadrado, que se caracteriza por una fase activa con picos cuyos máximos y mínimos locales son casi constantes y tienen un valor por encima de una fase silenciosa.

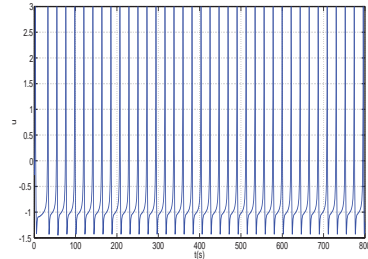
El beating se muestra en la Figura 2(b), en comparación con el bursting cuadrado carece de la fase silenciosa pero es igual en magnitud.

El bursting parabólico se muestra en la Figura 2(c) y tiene como característica principal un perfil que semeja una parábola. El bursting semiparabólico de la Figura 2(d) se caracteriza porque el valor de los mínimos locales en la fase activa está por debajo del valor en la fase silenciosa. Finalmente el tapered bursting tiene una fase activa sobreamortiguada y se muestra en la Figura 2(e).

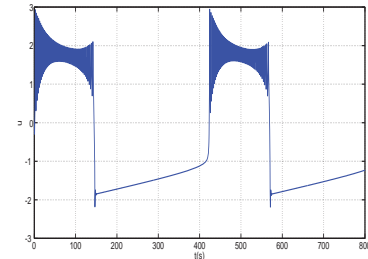
El valor de los parámetros  $a$ ,  $\eta$ ,  $\hat{v}$ ,  $\beta$  y  $u_\beta$  del modelo descrito en la Ecuación (1) determina el tipo de dinámica generado, que puede corresponder a alguno de los cinco tipos de bursting descritos. Para cada uno de ellos, el valor de los



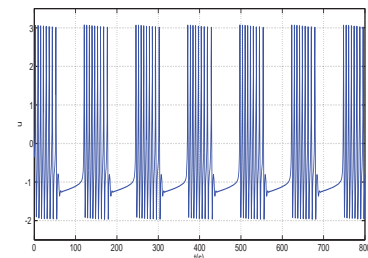
(a)



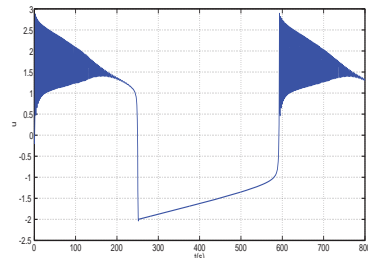
(b)



(c)



(d)



(e)

Figura 2. Dinámica de la variable de estado  $u$  del sistema dinámico de la célula beta cuando el bursting es del tipo (a) cuadrado, (b) beating, (c) parabólico, (d) semiparabólico y (e) tapered.

parámetros  $a$ ,  $\beta$  y  $u_\beta$  permanece fijo con los valores siguientes:

$$a = 1/4, \quad \beta = 4, \quad u_\beta = -0.954. \quad (3)$$

Los valores de los parámetros  $\eta$  y  $\hat{u}$  varían de acuerdo a la Tabla I para representar cada tipo de bursting.

Tabla I  
VALOR DE LOS PARÁMETROS DEL MODELO PARA CINCO TIPOS DE BURSTING.

Tipo de bursting	$\eta, \hat{u}$
Cuadrado	(3/4, 3/2)
Beating	(1, 3/2)
Parabólico	(3/4, 1)
Semiparabólico	(5/4, 1)
Tapered	(1/2, 9/4)

El bursting depende de escalas de tiempo distintas, por lo que una constante multiplica la tercera ecuación acoplada en el sistema de ecuaciones (1) y toma el valor de  $\epsilon = 0.0025$ .

Para cada tipo de bursting se ha obtenido la dinámica de los exponentes de Lyapunov. Para todos los tipos de bursting se tienen dos exponentes de Lyapunov negativos y uno igual a cero, lo que de acuerdo a la clasificación de Wolf [7], da como resultado un atractor del tipo ciclo límite. En la Tabla II se muestran los exponentes de Lyapunov obtenidos para cada tipo de bursting y el atractor resultante se muestra en la Figura 3.

Tabla II  
EXONENTES DE LYAPUNOV DE CADA TIPO DE BURSTING.

Tipo de bursting	Exponentes de Lyapunov
Cuadrado	0.04, -0.21, -1.22
Beating	0.04, -0.23, -1.09,
Parabólico	0.03, -0.44, -0.44
Semiparabólico	0.03, -0.31, -0.31
Tapered	0.05, -0.13, -2.46

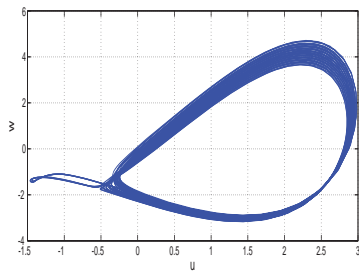


Figura 3. Diagrama de fase del sistema la célula beta produciendo bursting.

### III. SINCRONIZACIÓN CON RED COMPLEJA

Tomando en cuenta los cinco tipos de bursting analizados, se propone la sincronización de múltiples células beta exhibiendo diversos tipos de bursting por medio de una red compleja de tipo *mundo pequeño*.

Las redes complejas son modelos matemáticos diseñados para capturar características topológicas de sistemas complejos del mundo real [8]. El análisis de redes complejas se lleva a cabo a través de la teoría de grafos, que permite caracterizar dichos sistemas representando a cada individuo con un nodo que se encuentra interconectado con otros por medio de enlaces que emulan interacciones reales.

Para modelar con mayor precisión sistemas reales existen diversas topologías de entre las cuales la denominada mundo pequeño es la que mejor representa sistemas biológicos, ya que no es ni completamente regular ni completamente aleatoria. Las redes de mundo pequeño tienen un alto coeficiente de agrupamiento como una red regular y una distancia promedio corta tal como las redes aleatorias [5].

Una de las características más importantes de una red es la distribución de los enlaces  $P(k)$ , que es la probabilidad de que un nodo cualquiera tenga  $k$  conexiones. En el caso de la red de tipo mundo pequeño, esta distribución toma la forma de una ley de potencias de la Ecuación 4,

$$P(k) = Ck^{-\gamma}, \quad (4)$$

donde  $\gamma$  es real positivo.

#### III-A. Sincronización con redes complejas

Para una red de  $N$  nodos idénticos linealmente acoplados a través de la primera variable de estado donde cada nodo es un subsistema dinámico  $n$ -dimensional, las ecuaciones de estado son [6, 9]:

$$\begin{aligned} \dot{x}_{i1} &= f_1(\mathbf{x}_i) + c \sum_{j=1}^N a_{ij} x_{j1} \\ \dot{x}_{i2} &= f_2(\mathbf{x}_i) \\ &\vdots \\ \dot{x}_{in} &= f_n(\mathbf{x}_i), \end{aligned} \quad (5)$$

donde  $i = 1, 2, \dots, N$ ,  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{in}) \in \mathcal{R}^n$  son las variables de estado del nodo  $i$ ,  $c > 0$  es la fuerza de acoplamiento y  $A = (a_{ij})_{N \times N}$  es la matriz de acoplamiento.

La matriz de acoplamiento  $A$  es simétrica e irreducible y sus elementos  $a_{ij}$ , cuando ( $i \neq j$ ) solo pueden tomar el valor de 1 ó 0, de acuerdo a la condición siguiente:

- $a_{ij} = 1$  si los nodos  $i$  y  $j$  están unidos con un enlace.
- $a_{ij} = 0$  si los nodos  $i$  y  $j$  no están unidos.

Por su parte, los elementos de la diagonal son iguales al recíproco del grado de nodo, es decir:

$$a_{ii} = - \sum_{j=1, j \neq i}^N a_{ij}, \quad i = 1, 2, \dots, N.$$

#### III-B. Sincronización diversos tipos de bursting

Recientes estudios se han ocupado de las condiciones que podría afectar la sincronización del sistema del islote de Langerhans por medio de redes complejas de tipo mundo pequeño [3], en este trabajo se analiza la sincronización de

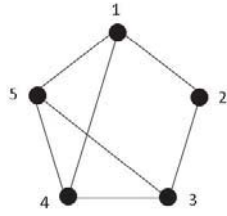


Figura 4. Red de mundo pequeño con cinco nodos.

células beta dentro del islote de Langerhans y se propone la red de cinco nodos de tipo mundo pequeño de la Figura 4.

Dentro de esta red, cada nodo se caracteriza con el modelo matemático de la Ecuación 1. Variando los valores de los parámetros de acuerdo a la Tabla I, es posible analizar la sincronización cuando los nodos dentro de la misma red exhiben distintos tipos de bursting. Debido a que el bursting cuadrado es el característico de la célula beta, se determinó realizar las pruebas de sincronización cuando una fracción de la red presenta bursting cuadrado y el resto de los nodos algún otro tipo de bursting.

Para la red de la Figura 4, la matriz de acoplamiento es:

$$\mathbf{A} = \begin{bmatrix} -3 & 1 & 0 & 1 & 1 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -3 & 1 & 1 \\ 1 & 0 & 1 & -3 & 1 \\ 1 & 0 & 1 & 1 & -3 \end{bmatrix} \quad (6)$$

#### IV. RESULTADOS DE SIMULACIÓN

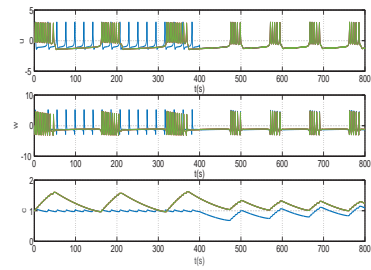
Se realizaron los siguientes cuatro tipos de pruebas para sincronización por medio de la red de la Figura 4:

1. Sincronización de bursting cuadrado con beating.
2. Sincronización de bursting cuadrado con parabólico.
3. Sincronización de bursting cuadrado con semiparabólico.
4. Sincronización de bursting cuadrado con tapered.

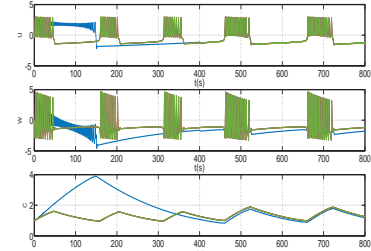
De los cinco nodos en la red, cuatro fueron caracterizados con bursting cuadrado y solo uno con otro tipo de bursting. A continuación se presentan los resultados obtenidos en cada caso. Como parte de las pruebas, se modificó la posición del nodo con un tipo de bursting distinto al cuadrado dentro de la red, pero en cada caso se observaron dinámicas resultantes similares, por lo que se concluye que para esta serie de pruebas, el grado de nodo no afecta la sincronización.

En la Figura 5 se muestran los resultados de la sincronización de diferentes tipos de bursting con un tiempo de simulación de 800s y acoplados a través de la primera variable de estado a partir de 400s. En todos los casos de sincronización presentados, las condiciones iniciales para cada sistema son las mismas, por lo tanto, solo se observan dos curvas en las gráficas.

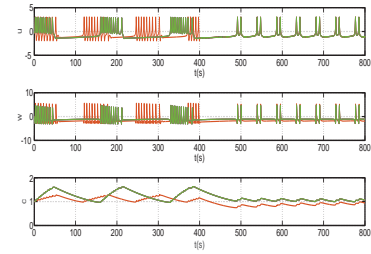
En la Figura 5(a) se muestran los resultados de la sincronización de cuatro células beta exhibiendo el bursting cuadrado



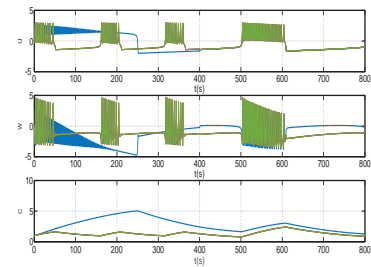
(a)



(b)



(c)



(d)

Figura 5. Sincronización de cuatro nodos exhibiendo bursting cuadrado con otro generando bursting de tipo: (a) beating, (b) parabólico, (c) semiparabólico y (d) tapered.

característico, con una célula que genera beating. Se observa que a partir de  $t = 400s$ , que es el momento en que se aplica el control, todas las curvas se sincronizan en una nueva dinámica distinta de las dos originales, caracterizada por una fase activa más corta que la del bursting cuadrado y un periodo más corto. La variable de estado  $c$  no llega a sincronizarse dentro del tiempo de simulación, pero el error va disminuyendo aunque

su amplitud se mantiene por debajo que la original.

Los resultados de la sincronización de cuatro células beta exhibiendo el bursting cuadrado característico, con una célula que genera bursting parabólico se muestran en la Figura 5(b). A partir del momento en que se aplica el control, en  $t = 400s$ , todas las curvas se sincronizan en una nueva dinámica distinta de las dos originales, caracterizada por una fase activa más larga que la del bursting cuadrado, al igual que una fase silenciosa más amplia y un periodo más largo. La variable de estado  $w$  no se sincroniza completamente durante la fase silenciosa.

En la Figura 5(c) se observan los resultados de la sincronización de cuatro células beta exhibiendo el bursting cuadrado característico y una que genera bursting semiparabólico. A partir de  $t = 400s$ , que es el momento en que se aplica el control, todas las curvas exhiben una nueva dinámica distinta de las originales. Las fases activa y silenciosa se hacen más cortas y el periodo disminuye aproximadamente a la mitad del observado en el bursting cuadrado. En este caso las variables de estado  $w$  y  $c$  no se sincronizan correctamente. Se observa que tanto en la fase silenciosa como la activa se mantiene un error constante en magnitud.

Finalmente los resultados de la sincronización de cuatro células beta exhibiendo el bursting cuadrado característico y una célula que genera tapered bursting se muestran en la Figura 5(d). A partir del momento en que se aplica el control, en  $t = 400s$ , todas las curvas se sincronizan en una nueva dinámica distinta de las dos originales, caracterizada por una fase activa más larga que la del bursting cuadrado, al igual que una fase silenciosa más amplia y un periodo más largo. La variable de estado  $w$  no se sincroniza completamente durante la fase silenciosa.

## V. CONCLUSIÓN

Se empleó el modelo matemático de Pernarowski para caracterizar las células beta. Mediante el diseño de una red compleja de tipo mundo pequeño se analizó la sincronización entre cinco distintos tipos de bursting. Se encontró que la sincronización de estas dinámicas da como resultado una nueva, distinta de las originales, con características intermedias, sin embargo, excepto por el caso de la sincronización de bursting cuadrado con bursting semiparabólico, se obtiene el bursting cuadrado característico de la célula beta.

En algunos casos, la sincronización de todas las variables de estado no se logra completamente ya que la acción de control se ha aplicado solo a la variable de estado  $u$ . En este caso pueden mejorarse los resultados aplicando también retroalimentación en una o las otras dos variables de estado.

El grado de nodo no afectó la sincronización en la serie de pruebas realizadas, lo que se corroboró cambiando la localización de los sistemas dentro de la red.

## AGRADECIMIENTOS

Esta investigación ha sido parcialmente soportada por el proyecto VIEP-BUAP 2016.

## REFERENCIAS

- [1] M. Pernarowski. *Fast and slow subsystems for a continuum model of bursting activity in the pancreatic islet*. SIAM Journal on Applied Mathematics, Vol 58, no. 5, pp 1667-1687, 1998.
- [2] M. Pernarowski, *Fast subsystem bifurcations in a slowly varying Lienard system exhibiting bursting*. SIAM Journal on Applied Mathematics, Vol 54, no.3, pp 814-832, 1994.
- [3] A.K. Barua, et al, *Isles within islets: The lattice origin of small-world networks in pancreatic tissues*. Physica D: Nonlinear Phenomena, Vol 315, pp 49-57, 2016.
- [4] R. Griffiths, "Return map characterizations of singular solutions for a model of bursting with two slow variables". Ph.D. Thesis, Montana State University, 2003.
- [5] D. J. Watts, et al, *Collective dynamics of a 'small world' networks*. Nature, Vol 393, no. 6684, pp 440-442, 1998.
- [6] X. F. Wang, et al, *Synchronization in small-world dynamical networks*. International Journal of Bifurcation and Chaos, Vol 12, no. 1, pp 187-192, 2002.
- [7] A. Wolf, et al, *Determining Lyapunov exponents from a time series*. Elsevier Physica D: Nonlinear Phenomena, Vol 16, no. 3, pp 285-317, 1985.
- [8] S. Boccaletti, et al, *Complex networks: Structure and dynamics*. Physics reports, Vol 424, no. 4, pp 175-308, 2006.
- [9] Z. Li, et al, *Global synchronization and asymptotic stability of complex dynamical networks*. Circuits and Systems II: Express Briefs, IEEE Transactions on, Vol 53, no. 1, pp 28-33, 2006.

# Algoritmo de rastreo aplicado al tratamiento de niños con autismo

Leopoldo N. Gaxiola, Victor H. Diaz-Ramirez, Andrés Calvillo-Téllez, Andres Cuevas  
 Instituto Politécnico Nacional - CITEDI  
 Ave. Instituto Politécnico Nacional 1310, Nueva Tijuana  
 Tijuana B.C. 22435, México  
 lgaxiola@citedi.mx

**Resumen**—En este trabajo se presenta un algoritmo de rastreo aplicado en apoyo en la medición del déficit de atención en niños con autismo. El primer paso es aplicar el algoritmo de rastreo en el rostro del niño para segmentarlo en cada cuadro de la escena. Posteriormente el algoritmo es aplicado nuevamente para rastrear la dirección del ojo y determinar el grado de atención del niño. El algoritmo es capaz de seguir el rostro y el ojo con invarianza a pose, gesticulaciones, oclusiones y fondo variado. Se muestran los resultados obtenidos en una secuencia de video de la vida real.

**Palabras claves**—Rastreo de rostros, rastreo del ojo, filtros de correlación.

## I. INTRODUCCIÓN

El autismo es un trastorno que puede presentarse desde el nacimiento o durante el desarrollo que afecta a los comportamientos humanos esenciales, como la interacción social, la capacidad de comunicación, la imaginación, y la interacción con los semejantes [1]. Por lo general, el autismo tiene efectos de larga duración sobre la capacidad para socializar de los niños, para cuidar de sí mismos, y de participar en la comunidad. El autismo es un trastorno del desarrollo cuyo origen es neurobiológico y se define sobre la base de características de comportamiento y de desarrollo.

Por medio de terapias se ha tratado de mejorar la calidad de vida de los niños que presentan este trastorno. Las actividades de las terapias son realizadas en ambientes aislados para evitar distractores externos. Al recibir terapia médica, los niños son grabados realizando diferentes actividades guiadas con el fin de que un especialista analice el progreso del niño. La tarea de analizar cada uno de los videos detenidamente para clasificar cada uno de los comportamientos es muy demandante tanto física como mentalmente. En apoyo a esta actividad en este trabajo se propone un sistema que ayude al especialista a analizar de manera automática las grabaciones de video de terapias de niños con autismo.

El rastreo de rostros consiste en la estimación de la trayectoria de un rostro mientras éste se mueve a través de una zona de detección [2], [3]. Las dificultades que presenta el rastreo de rostros son: la escena observada comúnmente esta degradada por ruido aditivo, presencia de un fondo variado, modificaciones geométricas del rostro a rastrear como pose y escala, e iluminación no uniforme.

Varios enfoques han sido propuestos para abordar el problema de rastreo de objetos [4], [5]. Algunos algoritmos requieren un sólido conocimiento *a-priori* del objetivo antes

de comenzar el rastreo [6]. Otros algoritmos sólo requieren un conocimiento mínimo del objetivo ya que el sistema de rastreo cuenta con mecanismos de aprendizaje y de adaptación en línea [5], [7]. Recientemente, se ha sugerido el uso de un filtro de correlación adaptativo para el rastreo de objetos en tiempo real [8]. Este método ha demostrado que es competitivo con los algoritmos de rastreo estándar, pero con menos complejidad [9], [5]. En este enfoque, un filtro de correlación (plantilla) se utiliza para detectar y localizar el objetivo dentro de la escena observada en cada cuadro. La plantilla se actualiza en línea (adaptativo) de acuerdo con la información extraída de observaciones de la escena (actual y anteriores), y teniendo en cuenta diferentes vistas del objetivo dentro de la escena. El filtro de correlación utilizado en este enfoque es el filtro de suma mínima de salida de error cuadrático (MOSSE, del inglés *Minimum Output Sum of Squared Error*) [8].

En el presente trabajo se propone un algoritmo para rastreo de objetos, aplicado tanto en el rostro y en los ojos de los niños. El algoritmo esta basado en un filtro de correlación compuesto localmente adaptativo sintetizado como una combinación de plantillas óptimas, diseñadas para la detección y estimación de la ubicación del objetivo dentro de la escena observada. Primeramente, se construye un conjunto de versiones geoméricamente distorsionadas de la imagen del rostro de referencia. Después, una plantilla (filtro óptimo) se diseña para cada imagen, y posteriormente, las plantillas se combinan para formar un filtro de correlación compuesto. Con el filtro construido se trata de localizar el rostro en la escena, si este no es capaz de localizarlo se ejecutan una etapa de reinicialización para tratar de redetectarlo.

Este artículo esta organizado de la siguiente forma. En la sección 2 se explica el diseño del filtro propuesto para el rastreo de rostros y ojos. La sección 3 describe el algoritmo propuesto para el rastreo de objetos. Los resultados obtenidos en simulaciones con nuestra propuesta son presentados y discutidos en la sección 4. Finalmente, la sección 5 presenta nuestras conclusiones.

## II. MARCO TEÓRICO

Se presenta el diseño de un filtro de correlación capaz de reconocer un rostro cuando se encuentra embebido en un fondo disjunto, y la escena esta degradada con ruido aditivo. Además, el filtro es capaz de reconocer versiones geoméricamente modificadas del rostro, como rotaciones y escalamientos. Es importante mencionar que el algoritmo se describe con su

aplicación en el rastreo de rostros solamente ya que es el mismo procedimiento para el rastreo de los ojos.

Sea  $T = \{t_i(x, y); i = 1, \dots, N\}$  un conjunto de imágenes que contiene versiones geoméricamente modificadas del rostro a reconocer. La escena de entrada se supone que está compuesta por el objeto  $t(x, y)$  embebido en un fondo disjunto  $b(x, y)$  en las coordenadas desconocidas  $(\tau_x, \tau_y)$ , y toda la escena esta degradada con ruido aditivo  $n(x, y)$ , como a continuación:

$$f(x, y) = t(x - \tau_x, y - \tau_y) + b(x, y) \bar{w}(x - \tau_x, y - \tau_y) + n(x, y), \quad (1)$$

donde  $\bar{w}(x, y)$  es una función binaria definida como cero dentro del área del rostro, y uno en cualquier otro lugar. Dentro del estado del arte del diseño de filtros de correlación para el reconocimiento de objetos, se sabe que el filtro óptimo para detectar un objeto a partir de la Ec. (1), en términos de la relación señal a ruido (SNR) [10], y la varianza mínima de las mediciones de errores de localización (LE), es el filtro GMF (del inglés, *Generalized Matched Filter*) [11], [12], cuya respuesta de frecuencia está dada por

$$H^*(u, v) = \frac{T(u, v) + \mu_b \bar{W}(u, v)}{P_b(u, v) \otimes |\bar{W}(u, v)|^2 + P_n(u, v)}. \quad (2)$$

En la Ec. (2),  $T(u, v)$  y  $\bar{W}(u, v)$  son la transformada de Fourier de  $t(x, y)$  y  $\bar{w}(x, y)$ , respectivamente;  $\mu_b$  es el valor promedio del fondo  $b(x, y)$ ; y  $P_b(u, v)$  y  $P_n(u, v)$  representan la densidad espectral de potencia de  $b_0(x, y) = b(x, y) - \mu_b$  y  $n(x, y)$ , respectivamente. El símbolo “ $\otimes$ ” denota convolución.

Sea  $h_i(x, y)$  la respuesta al impulso de un GMF diseñada para reconocer la  $i$ -ésima vista disponible del objeto  $t_i(x, y)$  en  $T$ . Sea  $H = \{h_i(x, y); i = 1, \dots, N\}$  el conjunto de todas las respuesta impulso generadas por el GMF para todas las imágenes de entrenamiento  $t_i(x, y)$ . Además, sea  $S = \{s_i(x, y); i = 1, \dots, M\}$  un conjunto de imágenes que contiene  $M$  patrones a ser rechazados. Los patrones a ser rechazados, pueden ser fragmentos del fondo o cualquier otro objeto falso. La idea es sintetizar un filtro capaz de reconocer todas las vistas de  $T$  y rechazar todos los patrones falsos de  $S$ , mediante la combinación de las plantillas óptimas contenidas en  $H$ , y mediante el uso de solo una operación de correlación. El filtro requerido  $p(x, y)$ , se construye como a continuación:

$$p(x, y) = \sum_{i=1}^N \alpha_i h_i(x, y) + \sum_{i=N+1}^{N+M} \alpha_i s_i(x, y), \quad (3)$$

donde los coeficientes  $\{\alpha_i; i = 1, \dots, N + M\}$  son elegidos para satisfacer los valores de salida predefinidos para cada patrón en  $U = T \cup S$ . Sea  $\mathbf{R}$  una matriz con  $N + M$  columnas, donde cada columna es la versión vector de cada elemento de  $U$ , y sea  $\mathbf{a} = [\alpha_i; i = 1, \dots, N + M]^T$  un vector de coeficientes incógnita. De esta manera, la Ec. (3) se puede reescribir como

$$\mathbf{p} = \mathbf{R}\mathbf{a}. \quad (4)$$

Ahora, sea  $\mathbf{u}$  el vector de restricciones dado por

$$\mathbf{u} = \left[ \underbrace{1, \dots, 1}_{N \text{ unos}}, \underbrace{0, \dots, 0}_{M \text{ ceros}} \right]^T. \quad (5)$$

El vector  $\mathbf{u}$  contiene las respuestas deseadas para los patrones de entrenamiento. Sea  $\mathbf{Q}$  una matriz cuyas columnas son los elementos de  $U$ . Las restricciones de salida pueden expresarse como

$$\mathbf{u} = \mathbf{Q}^+ \mathbf{p}, \quad (6)$$

donde el superíndice  $+$  denota complejo conjugado. Al sustituir la Ec. (4) dentro de la Ec. (6), se obtiene

$$\mathbf{u} = \mathbf{Q}^+ \mathbf{R}\mathbf{a}. \quad (7)$$

Por lo tanto, la solución para  $\mathbf{a}$ , es

$$\mathbf{a} = [\mathbf{Q}^+ \mathbf{R}]^{-1} \mathbf{u}. \quad (8)$$

Finalmente, al sustituir la Ec. (8) en la Ec. (4), la solución del filtro compuesto esta dada por

$$\mathbf{p} = \mathbf{R} [\mathbf{Q}^+ \mathbf{R}]^{-1} \mathbf{u}. \quad (9)$$

Cabe mencionar que el valor del pico de correlación cuando se utiliza el filtro dado en la Ec. (9), se espera que sea cercano a la unidad para los objetos de la clase verdadera, y cercano a cero para los objetos de la clase falsa.

### III. IMPLEMENTACIÓN

En esta sección se describe el algoritmo propuesto para el rastreo de rostros y ojos basado en un filtro de correlación adaptativo. El algoritmo es robusto a cambios de escala, pose y manejo de oclusiones. El algoritmo comienza con una etapa de inicialización en donde se define el rostro a ser rastreado. A continuación, el algoritmo genera varias versiones geoméricamente modificadas de la imagen de referencia del rostro y se construyen los conjuntos  $T$  y  $H$ . Además, si se tiene información *a-priori* de los patrones no deseados a ser rechazados, como otros objetos o fragmentos de fondo, podemos utilizarlos para construir el conjunto  $S$ . Luego, un filtro compuesto es sintetizado por medio de la Ec. (9). El siguiente paso es correlacionar el filtro compuesto con una región de interés (fragmento de un nuevo cuadro de la escena), y calcular la capacidad de discriminación (DC) del plano de intensidad de correlación resultante. La DC es una medida de la habilidad del filtro para distinguir un objetivo de los objetos no deseados; la DC se define por [11]

$$DC = 1 - \frac{|c^b|^2}{|c^t|^2}, \quad (10)$$

donde  $c^b$  es el valor del lóbulo de correlación máximo en la zona del fondo y  $c^t$  es el valor del pico de correlación generado por el rostro. Un valor DC cercano a la unidad indica que el filtro tiene una buena capacidad para distinguir entre el objetivo y cualquier objeto falso. Los valores negativos de DC indican que el filtro no es capaz de detectar al rostro objetivo. A continuación, si la DC obtenida es mayor que un umbral predefinido ( $DC > DC_{th}$ ), entonces el rostro objetivo se considera como detectado. Luego en base a la estimación de la posición generada se predice la siguiente posición y

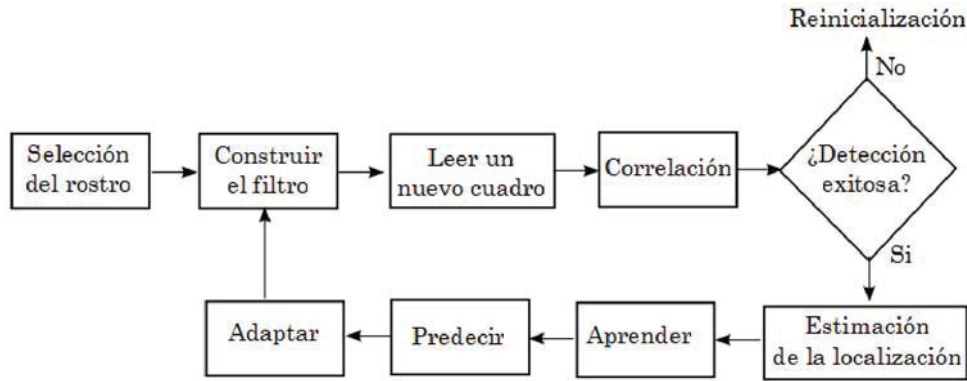


Fig. 1: Diagrama a bloques del sistema propuesto para el rastreo de rostros.

se adapta el sistema en una región de interés de la escena alrededor de dicha posición. Finalmente, el filtro compuesto se actualiza con una nueva imagen de referencia con el objetivo detectado en el cuadro actual. De lo contrario, el objetivo es rechazado. Se puede observar, que el filtro compuesto se actualiza en línea para cada cuadro sin la necesidad de algún proceso de entrenamiento supervisado (fuera de línea) [6]. El funcionamiento del algoritmo se representa en la Fig. 1, y sus pasos se resumen en el algoritmo 1.

---

**Algoritmo 1** Pseudocódigo del sistema de rastreo de rostros.

- PASO 1 Leer la escena de entrada  $f_k(x, y)$  de la secuencia de video y seleccionar el rostro a rastrear con una pequeña ventana  $b_k(x, y)$
- PASO 2 Crear versiones modificadas geoméricamente de  $b_k(x, y)$  y construir los conjuntos  $T$  y  $H$ .
- PASO 3 Construir un filtro compuesto  $p_k(x, y)$  (con Ec. (9)) entrenado con los conjuntos  $T$  y  $H$  y construir un filtro adaptativo de la siguiente manera:
- $$p_k(x, y) = \alpha p_k(x, y) + (1 - \alpha) p_{k-1}(x, y), \quad (11)$$
- donde  $\alpha$  es una constante y  $p_{k-1}(x, y)$  es el filtro para el cuadro de la escena anterior. Nota:  $p_{k-1}(x, y) = 0$  cuando  $k = 1$ .
- PASO 4 Calcular la función de correlación cruzada  $c_k(x, y)$  entre el filtro adaptativo  $p_k(x, y)$  y la ventana  $b_k(x, y)$ . Después calcular la DC por medio de la Ec. (10).
- PASO 5 Si  $DC > DC_{th}$ , estimar la posición actual del rostro como  $(\tau_x, \tau_y) = \operatorname{argmax}_{x,y} \left\{ |c_k(x, y)|^2 \right\}$ , predecir la nueva posición del rostro para el siguiente cuadro y enfocar el algoritmo en una área más pequeña. De otra manera, leer un nuevo cuadro de la escena y actualizar la ventana  $b_k(x, y)$  mediante el desplazamiento de su origen a la estimación de la posición actual del rostro. Ir al paso 2.
- 

1) *Predicción*: La predicción de la posición del rostro en el siguiente cuadro de la escena se lleva a cabo por medio series de tiempo. Esta forma de predicción, representa un método adaptativo que aprende la dinámica del objetivo por medio de las observaciones pasadas, por lo cual se puede ajustar a

diferentes cambios en la dinámica del objeto. La pérdida se inicia a partir de las primeras cinco posiciones estimadas del objetivo, realizadas a través del algoritmo propuesto. Las series de tiempo utilizadas son de segundo orden y se utilizan dos series independientes, una para la predicción del eje  $x$  y otra para la del eje  $y$ . Estas series, se describen matemáticamente como [13]

$$\hat{x}_{k+1} = a_x + b_x(k+1) + c_x(k+1)^2 \quad (12)$$

y

$$\hat{y}_{k+1} = a_y + b_y(k+1) + c_y(k+1)^2, \quad (13)$$

respectivamente, donde  $k$  es el índice del cuadro actual y  $k+1$  es el índice del cuadro siguiente,  $\hat{x}_{k+1}$  es el valor de la predicción para el cuadro siguiente en el eje  $x$  y  $\hat{y}_{k+1}$  el valor en el eje  $y$ , respectivamente.

Los coeficientes  $a$ ,  $b$  y  $c$  para las Ecs. (12) y (13), se calculan de la siguiente manera [13]

$$\sum_{k=1}^n x_k = an + c \sum_{k=1}^n k^2, \quad (14)$$

$$\sum_{k=1}^n k^2 x_k = a \sum_{k=1}^n k^2 + c \sum_{k=1}^n k^4 \quad (15)$$

y

$$b = \frac{\sum_{k=1}^n k x_k}{\sum_{k=1}^n k^2}, \quad (16)$$

donde  $a$ ,  $b$  y  $c$  son los valores que se desea encontrar,  $x$  son los valores en el eje  $x$  de las posiciones anteriores  $k$  es el número del cuadro y  $n$  define el tamaño de la muestra de las posiciones anteriores (cinco para este caso). Las Ecs. (14-16) están definidas para encontrar los valores para el eje  $x$  y es el mismo procedimiento para el caso del eje  $y$ .



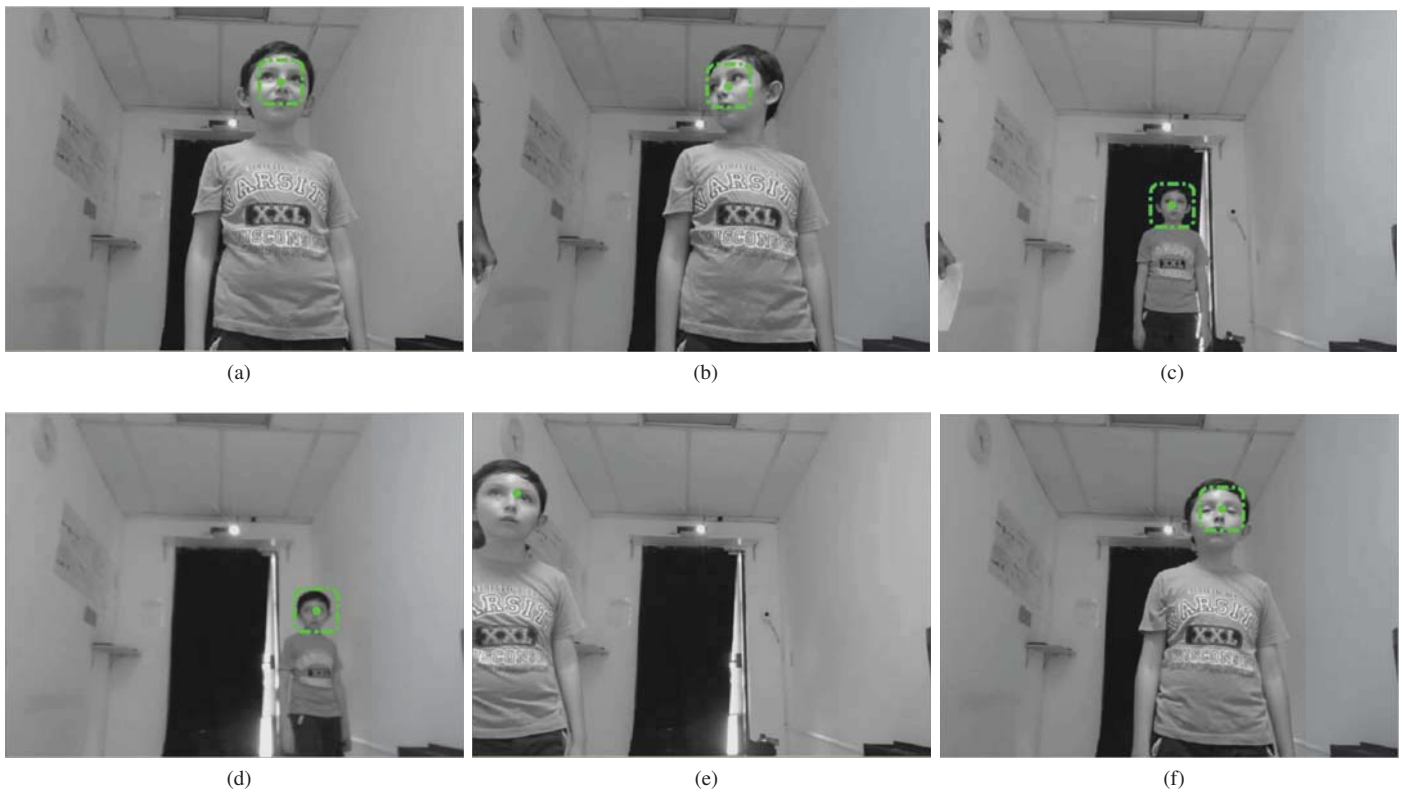


Fig. 2: Ejemplos de cuadros de la secuencia de video de prueba. (a) Objetivo (imagen de referencia), (b) cambios de pose, (c) cambio de escala, (d) desenfoque, (e) cerca del borde, (f) gesticulaciones.

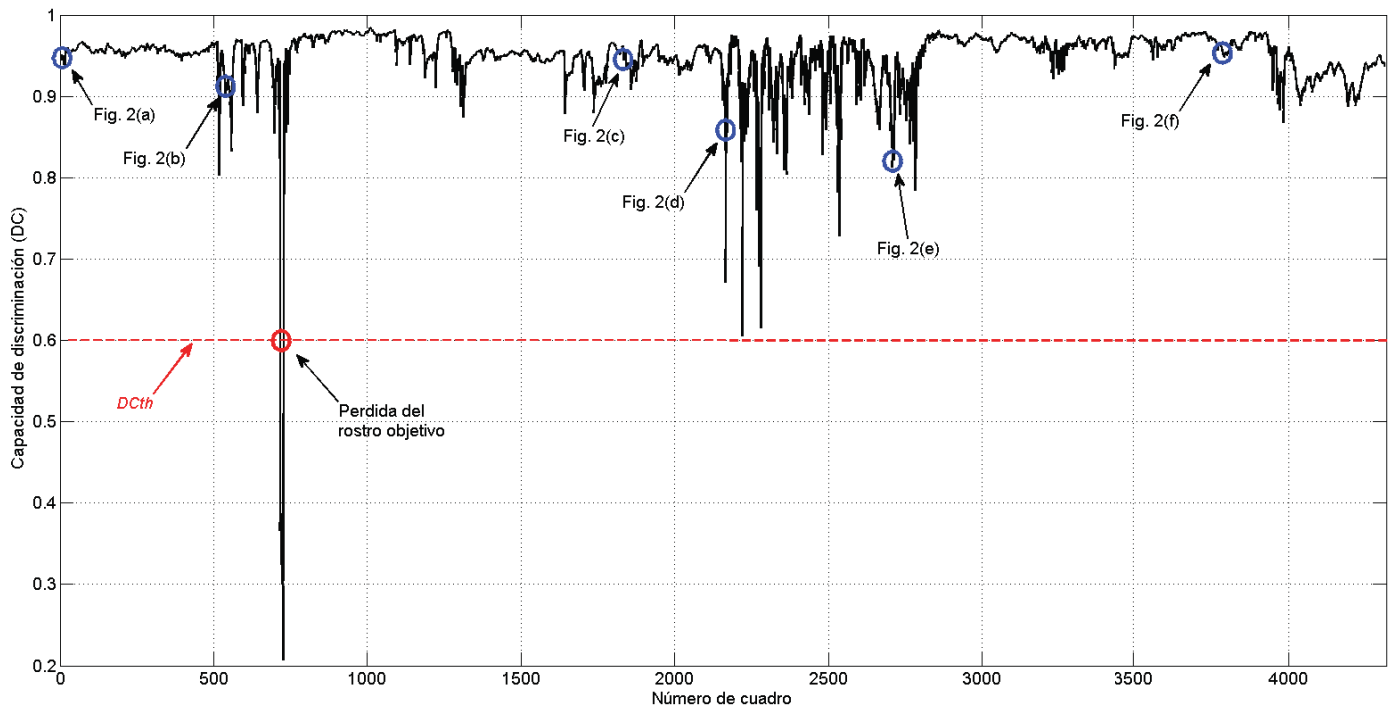


Fig. 3: Desempeño del algoritmo de rastreo en trminos de la DC al procesar 4315 cuadros de la escena.

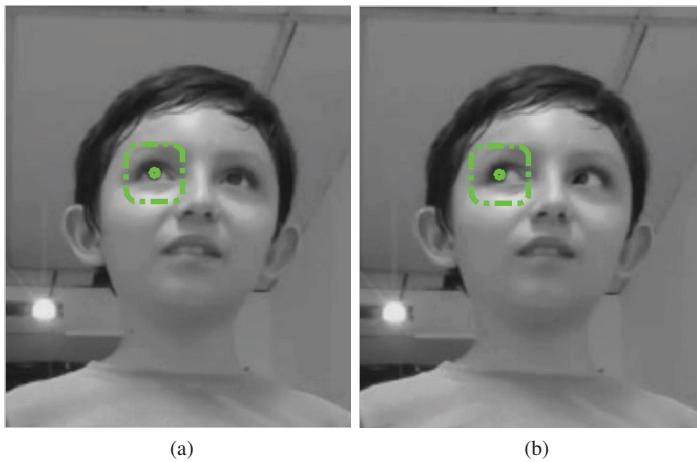


Fig. 4: Ejemplos de cuadros de la secuencia de video en el rastreo del ojo.

#### IV. RESULTADOS

El algoritmo fue probado utilizando una secuencia de video de la vida real que contiene cambios de pose, escala, desenfoque, distorsión y oclusiones. El algoritmo se evaluó en términos de eficiencia de detección y precisión en el rastreo. La eficiencia en la detección se caracteriza por la DC. La precisión de seguimiento se evalúa en términos de errores de rastreo. La secuencia de prueba contiene 4315 cuadros de escena monocromáticos, cada uno de ellos con  $1920 \times 1080$  píxeles. El rango de la señal es  $[0, 1]$  con 256 niveles de cuantificación. El algoritmo fue implementado en MATLAB en una computadora de escritorio con un procesador Intel i7 3,4 GHz con 16 GB de memoria RAM, que se ejecuta en el sistema operativo Windows 7 de 64 bits.

La Fig. 2 muestra ejemplos de cuadros de la secuencia de prueba. El cuadrado verde identifica la posición estimada del rostro obtenida con el algoritmo propuesto. Observe que el algoritmo propuesto es capaz de realizar el seguimiento del objetivo en todos los cuadros, incluso en presencia de cambios de pose, escalamiento, desenfoque y oclusiones. Es importante mencionar que el algoritmo propuesto ha sido probado con éxito en diferentes videos de la vida cotidiana, en donde se contemplan perturbaciones como cambios de iluminación, oclusiones parciales, deformaciones, rotación y escala, y ruido [14].

La Fig. 3 muestra el valor de la DC obtenido por el algoritmo propuesto para todos los cuadros de la secuencia de prueba. Los círculos azules corresponden a los valores de DC de los rastreadores para los cuadros de la escena que se muestran en la Fig. 2. Se puede observar que el rastreador propuesto obtiene buen desempeño en términos de la DC. Además, la Fig. 3 muestra los cuadros de la escena en donde el algoritmo de rastreo pierde el objetivo. Esto se identifica por una elipse roja. Se observa que el algoritmo propuesto es capaz de detectar y seguir el rostro objetivo en un 99% de los cuadros de la secuencia de video. Además la Fig. 4 muestra ejemplos de cuadros de video del rastreo del ojo a partir de las imágenes del rostro generadas por el proceso anterior.

#### V. CONCLUSIONES

Se presentó un algoritmo para el rastreo de rostros y ojos en apoyo a terapias médicas de niños con autismo. El algoritmo detecta y localiza el rostro del paciente en cada cuadro de una secuencia de video utilizando un filtro de correlación compuesto que se adapta en línea. El filtro está construido por la combinación de plantillas óptimas diseñadas para la detección y la estimación de la posición del rostro. El filtro propuesto se implementó en el rastreo de rostros en una secuencia de video de la vida real. Los resultados de la simulación mostraron que el filtro propuesto es significativamente robusto y preciso, en presencia de cambios de pose, desenfoque, cambios de escala y oclusiones.

#### REFERENCIAS

- [1] C. Lord y J. P. McGee, *Educating children with autism*. National Academy Press Washington, DC, 2001.
- [2] C. Huang, X. Ding, y C. Fang, "Pose robust face tracking by combining view-based AAMs and temporal filters," *Comput. Visi. Image Und.*, vol. 116, no. 7, pp. 777–792, 2012.
- [3] M. Kim, S. Kumar, V. Pavlovic, y H. Rowley, "Face tracking and recognition with visual constraints in real-world videos," in *IEEE Conference on Computer Vision and Pattern Recognition, CVPR*, 2008, pp. 1–8.
- [4] A. Yilmaz, O. Javed, y M. Shah, "Object tracking: A survey," *ACM computing surveys (CSUR)*, vol. 38, no. 4, p. 13, 2006.
- [5] Y. Wu, J. Lim, y M.-H. Yang, "Online object tracking: A benchmark," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2013, pp. 2411–2418.
- [6] V. H. Diaz-Ramirez, K. Picos, y V. Kober, "Target tracking in nonuniform illumination conditions using locally adaptive correlation filters," *Opt. Commun.*, vol. 323, pp. 32–43, 2014.
- [7] Z. Kalal, K. Mikolajczyk, y J. Matas, "Tracking-learning-detection," *IEEE Trans. Pattern Anal. and Mach. Intell.*, vol. 34, no. 7, pp. 1409–1422, 2012.
- [8] D. S. Bolme, J. R. Beveridge, B. A. Draper, y Y. M. Lui, "Visual object tracking using adaptive correlation filters," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010, pp. 2544–2550.
- [9] J. F. Henriques, R. Caseiro, P. Martins, y J. Batista, "Exploiting the circulant structure of tracking-by-detection with kernels," in *Computer Vision—ECCV 2012*. Springer, 2012, pp. 702–715.
- [10] B. V. Kumar y L. Hasebrook, "Performance measures for correlation filters," *Appl. Opt.*, vol. 29, no. 20, pp. 2997–3006, 1990.
- [11] V. Kober y J. Campos, "Accuracy of location measurement of a noisy target in a nonoverlapping background," *JOSA A*, vol. 13, no. 8, pp. 1653–1666, 1996.
- [12] B. Javidi y J. Wang, "Design of filters to detect a noisy target in nonoverlapping background noise," *JOSA A*, vol. 11, no. 10, pp. 2604–2612, 1994.
- [13] D. C. Montgomery, C. L. Jennings, y M. Kulahci, *Introduction to time series analysis and forecasting*. John Wiley & Sons, 2015.
- [14] L. N. Gaxiola, V. H. Diaz-Ramirez, J. J. Tapia, y P. García-Martínez, "Target tracking with dynamically adaptive correlation," *Opt. Commun.*, vol. 365, pp. 140–149, 2016.

# Reconocimiento de armas de fuego tipo pistola utilizando visión artificial

Palacios Alvarado, Christian Alejandro  
Maestría en Sistemas  
Computacionales  
ITLP  
La Paz, B.C.S, México  
palaciosc12@gmail.com

Martínez Díaz, Saúl  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S, México  
saulmd@itlp.edu.mx

**Resumen**—En este artículo se presenta la implementación de un algoritmo para el reconocimiento de armas de fuego tipo pistola utilizando visión artificial. La visión artificial es una materia compleja que trata de imitar la capacidad que tiene el ser humano de interpretar una escena o imagen. Por ejemplo, la visión humana, es una tarea de procesamiento de información. Es la función del sistema visual para proporcionar una representación del entorno que nos rodea y, más importante, para suministrar la información que necesitamos para interactuar con el medio ambiente. Para ser específico, el cerebro lleva a cabo rutinariamente tareas de reconocimiento de percepción (por ejemplo, reconociendo una cara familiar incrustado en una escena desconocida) en aproximadamente 100-200 ms, mientras que las tareas de mucha menor complejidad toman mucho más tiempo en una potente computadora [1]. En la visión artificial existe una amplia gama de herramientas y métodos para poder lograr esto.

Es este trabajo se propone el reconocimiento de armas utilizando redes neuronales. Lo primero es combinar capas para obtener una imagen en escala de grises. Posteriormente se aplica lo que es la segmentación para eliminar el fondo de la imagen, esto con ayuda de los operadores morfológicos. Después se extraen los rasgos característicos con los invariantes de Hu y se aplican a la red para obtener su clasificación.

**Palabras claves**—Procesamiento de imágenes, red neuronal, Matlab, invariantes de Hu y operadores morfológicos.

## I. INTRODUCCIÓN

La inseguridad en México es uno de los principales problemas sociales que afectan al país. El Instituto Ciudadano de Estudios Sobre la Inseguridad (ICESI) ha reportado que el asalto a transeúnte sigue siendo el de mayor incidencia [2] y, en la mayoría de los casos, sin pistas para capturar a los delincuentes. La Figura 1 clasifica los diferentes tipos de delito y muestra su incidencia. En muchos lugares del país se ha optado por instalar cámaras de video vigilancia que almacenan los videos por cierto tiempo. Esta tecnología que actualmente se utiliza, necesita de personal que analice las escenas y dicte un juicio de las actividades que aparecen. Desafortunadamente es prácticamente imposible que una persona pueda analizar de forma simultánea escenas captadas con cámaras ubicadas en distintos lugares. Esto provoca que no se detecten muchas de las problemáticas documentadas en los videos.

ENTIDAD	ENSI-3	ENSI-5	ENSI-6	ENSI-7	ENTIDAD	ENSI-3	ENSI-5	ENSI-6	ENSI-7
Distrito Federal	19,700	25,700	22,800	26,500	Coahuila	14,000	4,900	15,900	8,200
Aguascalientes	12,300	6,700	15,900	19,800	Morelos	11,300	8,200	8,900	8,200
Quintana Roo	16,400	12,100	12,000	15,900	Campeche	9,200	7,500	6,300	7,700
México	16,100	15,100	15,600	15,800	Oaxaca	5,200	4,400	11,100	7,500
Baja California	20,700	13,600	15,900	13,300	Colima	6,300	4,300	12,300	7,500
Sonora	14,500	9,900	20,400	12,700	Zacatecas	4,300	2,100	6,300	7,200
Baja California Sur	10,800	6,700	15,100	11,900	Tabasco	6,200	6,800	ND	6,900
Chihuahua	10,500	12,600	15,300	11,700	Nayarit	5,500	5,600	10,800	6,600
Michoacán	9,400	6,400	12,700	11,000	Oaxaca	6,100	3,900	5,400	6,400
Guanajuato	7,700	11,400	11,000	10,800	Tlaxcala	5,900	5,600	6,100	5,900
Jalisco	14,500	10,800	14,100	10,500	Sinaloa	14,900	4,400	7,500	5,800
Nuevo León	8,000	12,500	12,600	10,200	Veracruz	4,100	6,900	4,400	5,700
San Luis Potosí	4,300	6,700	9,400	9,500	Tamaulipas	10,700	13,800	ND	5,400
Yucatán	14,700	4,400	5,100	9,200	Hidalgo	5,700	3,100	9,600	4,000
Durango	10,600	4,600	8,000	8,800	Guerrero	8,200	6,100	6,400	3,900
Puebla	11,800	10,900	7,700	8,500	Chiapas	1,800	3,100	3,700	2,400
					NACIONAL	11,200	10,500	12,000	11,000

Figura 1. Tabla de delitos por cada 100 habitantes.

## II. TRABAJOS PREVIOS RELACIONADOS

Hoy en día existen varios trabajos en el campo del reconocimiento de imágenes, cada uno de estos trabajos utiliza técnicas y algoritmos diferentes los cuales están acorde al problema que se intenta solucionar, así también utilizan diferentes tipos de herramientas para poder realizar el procesamiento de dichos algoritmos, como el uso de GPU's o FPGA's. Algunos trabajos donde encontramos el uso de algoritmos para el reconocimiento de imágenes, son el trabajo Automated Recognition of Firearms in Surveillance Video [3], en el que se habla de una algoritmo capaz de reconocer armas de fuego dentro de una escena, donde los algoritmos usados fueron implementados en un CPU, este trabajo es el que más se asemeja a la investigación presentada en este artículo. En el trabajo de Procesamiento de imágenes mediante operadores de morfología matemática implementados en un FPGA [4], se presentó la implementación de dichos operadores en un FPGA. La morfología matemática es una técnica de procesamiento de imágenes que se utilizó para erosionar o dilatar imágenes. El cálculo de estos operadores con lleva un alto costo computacional por eso es que se utilizó un FPGA para realizar el procesamiento más rápido de lo que lo realizaría en

una computadora convencional. Este trabajo nos servirá como referencia en un futuro para implementar dichos operadores dentro de esta tecnología, dado que en la investigación se utilizó la morfología matemática para procesar las imágenes. Así también está el trabajo de Reconocimiento y localización visual 3D de objetos para robots manipuladores [5], en el cual encontramos la implementación de los invariantes de Hu para discriminar la forma de un objeto.

### III. HERRAMIENTAS UTILIZADAS EN EL TRABAJO

**Matlab:** Matlab es un software matemático el cual nos ofrece una amplia gama de herramientas para el procesamiento de imágenes, entre otras. Éste se utilizó para programar los algoritmos implementados en el trabajo.

**Arma de fuego:** Se utilizó un arma de fuego (Pistola de gotcha de Airsoft), la cual tiene 18 cm. de longitud y 14 cm. de altura. A pesar de ser un arma de juguete cuenta con las dimensiones reales de una pistola Walther y cuenta con la misma apariencia, tal y como se muestra en la la figura 2.

**Banco de imágenes:** Se realizó un banco de 43 imágenes, dentro de todas se incluyó un arma de fuego para poder detectarla dentro de la misma. Junto a este objeto se colocaron otros del mismo color o diferente, diferentes volúmenes y tamaños así también, el arma fue girada en diferentes direcciones y ángulos. Las imágenes cuentan con una resolución de 640 x 480 pixeles en formato BMP.



Figura 2. Arma Walther de Airsoft.

**Red neuronal:** Las redes neuronales se caracterizan por tener la capacidad de organizar sus componentes estructurales, conocidos como neuronas, con el fin de realizar ciertos cálculos (por ejemplo, el reconocimiento de patrones, la percepción y el control de motores) muchas veces mas rápido que el ordenador digital más rapido actualmente [1]. En este trabajo se utilizó una red de tipo feed forward, la cual se programa en la aplicación Matlab. La red cuenta con 4 neuronas de entrada, 10 en la capa oculta y una de salida. La función de entrenamiento utilizada en la red neuronal fue la trainlm. En el siguiente diagrama se ve la estructura de la red neuronal utilizada.

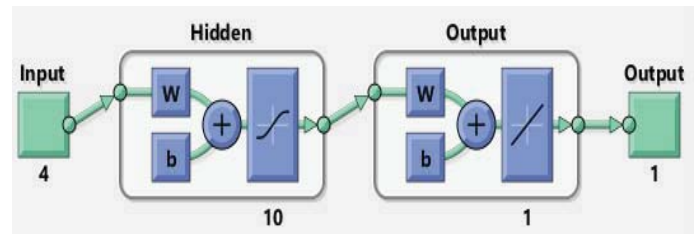


Figura 3. Estructura de la red neuronal.

### IV. DESCRIPCIÓN DEL ALGORITMO

En la figura 3 se muestra el algoritmo propuesto para la detección de armas, en los siguientes párrafos se describe los pasos del algoritmo.

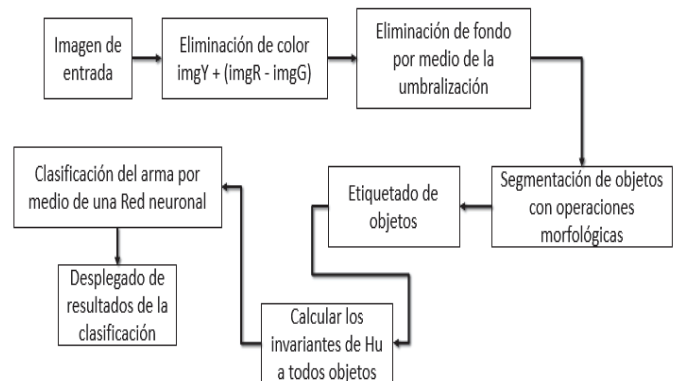


Figura 4. Diagrama de Activiad.

**Imagen de entrada:** En este proceso se toma una imagen del banco que se tiene, en ésta se encuentra el arma junto a otros objetos del mismo color o diferente, así también estos cuentan con diferente tamaño, tal y como se muestra en la figura 4.



Figura 5. Imagen a color del arma.

**Eliminación de color:** En este paso se procesa la imagen para eliminar el color y tratarla a escala de grises. Primero obtenemos la capa de iluminación de la imagen. Para hacer esto se utiliza la fórmula 1, la cual se define en [6].

$$imgY = ((2 * capaR) / 7) + ((4 * capaV) / 7) + (capaA / 7) \quad (1)$$

Una vez que se obtiene la iluminosidad se toman las capas roja y verde, con estas dos se relizó la fórmula 2.

$$O1 = imgR - imgG \quad (2)$$

Al resultado se le suma la capa de iluminación ( $imgY$ ), por lo que nos queda la fórmula 3.

$$imgTrabajar = O1 + imgY \quad (3)$$

**Eliminación de fondo:** Para este paso se utiliza una imagen en escala de grises, tal y como se muestra en la figura 5. Se utilizó la técnica de umbralización para de esta forma eliminar los objetos más claros dentro de la imagen (El arma dentro de la escena es de color negro), se optó por dejar los colores más cercanos al negro. A continuación se muestra el algoritmo 1, para obtener un umbral óptimo. Una vez que se tiene este umbral se utiliza para segmentar la imagen y hacerla binaria, tal y como se muestra en el algoritmo 2. En la figura 6 se muestra el resultado obtenido.

---

**Algoritmo 1** Pseudocódigo para obtener el umbral óptimo

---

```

while ( $T_{actual} - T_{vieja}$ ) >  $varDetener$  do
  for  $x \leftarrow 0; i < imgLongitudX; x \leftarrow x + 1$  do
    for  $y \leftarrow 0; y < imgLongitudY; y \leftarrow y + 1$  do
      if  $img(x, y) > T_{actual}$  then
        if  $img(x, y) > 0$  then
           $C1 = C1 + 1;$ 
           $G1 = G1 + img(x, y);$ 
        end if
      else
        if  $img(x, y) > 0$  then
           $C2 = C2 + 1;$ 
           $G2 = G2 + img(x, y);$ 
        end if
      end if
    end for
  end for
   $promedioG1 = G1/C1; promedioG2 = G2/C2;$ 
   $T_{vieja} = T_{actual};$ 
   $T_{actual} = (promedioG1 + promedioG2)/2;$ 
end while

```

---

**Algoritmo 2** Pseudocódigo de segmentación por umbral

---

```

for  $x \leftarrow 0; i < imgLongitudX; x \leftarrow x + 1$  do
  for  $y \leftarrow 0; y < imgLongitudY; y \leftarrow y + 1$  do
    if  $img(x, y) > T_{actual}$  then
       $imgSegmentada(x, y) = 0;$ 
    else
       $imgSegmentada(x, y) = 1;$ 
    end if
  end for
end for

```

---



Figura 6. Imagen en escala de grises.



Figura 7. Eliminación del fondo de la imagen.

**Segmentación de objetos:** Una vez que se tiene la imagen binaria, por haber separado el fondo. Se procede a aplicar las operaciones morfológicas para reducir el ruido que queda en la imagen. Primero se aplica una dilatación con un elemento estructurante de disco para rellenar algunos huecos que deja la eliminación de fondo. Este efecto se da debido a que la luz reflejada en el arma aclara estos huecos y con la umbralización se terminan desapareciendo. Después se aplicó una erosión para terminar de separar aquellos objetos o formas que aun se sobrepone. En la figura 7 se muestra el resultado de haber aplicado estos operadores.



Figura 8. Imagen dilatada y erosionada.

**Etiquetar objetos:** Se utiliza la función *bwlabel* de Matlab, para marcar los objetos dentro de la misma. La función *bwlabel* utilizó el criterio de 4 vecinos para etiquetar los objetos, una vez que todos los objetos estén etiquetados dentro de la imagen, se realizó el cálculo de los invariantes de Hu a todos los objetos etiquetados en la imagen.

**Cálculo de invariantes de Hu:** Los momentos geométricos cartesianos, han sido ampliamente usados en el análisis de formas, el reconocimiento de patrones y el análisis de texturas [7]. Se utilizan los invariantes a traslación, rotaciones y cambios de escala, dado que las imágenes que se tomaron al arma se encuentra rotada en varios ángulos y en diferente escala según la distancia de ésta y la cámara. A continuación se muestran las fórmulas para el cálculo de estos momentos.

$$\begin{aligned} \phi_1 &= \eta_{20} + \eta_{02} \\ \phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \phi_3 &= (\eta_{30} - 3\eta_{12})^2 + 3(\eta_{21} - \eta_{03})^2 \\ \phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ & \quad (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})] + 4\eta_{11}(\eta_{30} + \\ & \quad \eta_{12})(\eta_{21} + \eta_{03}) \\ \phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ & \quad (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned}$$

Se realizó el cálculo de los invariantes a todas las imágenes del banco con el que se cuenta. En la tabla 1 se muestra el mínimo y máximo de estos invariantes de Hu, que se obtuvieron al realizarle este cálculo al objeto en forma de arma de fuego que se localizaba en la imagen. Estos momentos están normalizados.

**Clasificación de objetos:** Los invariantes de Hu utilizados para la clasificación en la red neuronal fueron obtenidos después de haber realizado el cálculo a todas las imágenes del banco. Cabe mencionar que solo se están tomando los primeros 4 invariantes, debido a que después del número 5 estos no contribuyen en la discriminación de la forma. Para la clasificación de los objetos obtenidos en la imagen.

Tabla I  
MOMENTOS CENTRALES NORMALIZADOS DE ORDEN 3

	$\phi_1$	$\phi_2$	$\phi_3$	$\phi_4$	$\phi_5$	$\phi_6$	$\phi_7$
Mín	0.2402	0.0001	0.0020	0.0003	0.0000	0.0000	0.0000
Máx	0.3052	0.0311	0.0177	0.0019	0.0000	0.0001	0.0000

Se utilizó una red neuronal de tipo feed forward, la cual se describió anteriormente. La red neuronal procesa una matriz de 4x993. Esta matriz tiene los 4 invariantes de Hu de 993 objetos. Esta cantidad de objetos se obtuvieron al procesar todas las imágenes en el banco. En las gráficas siguientes se muestran el resultado del entrenamiento, validación y prueba, que arrojó la red neuronal. El valor de 0 representa los objetos que no son arma y el valor 1 representa el objeto en forma de arma. En los datos de entrenamiento se obtuvo un resultado de 0.95275 de efectividad, mientras que en los datos de validación se obtuvo una respuesta de 0.92239 y por último en los datos de prueba el resultado obtenido fue de 0.94044.

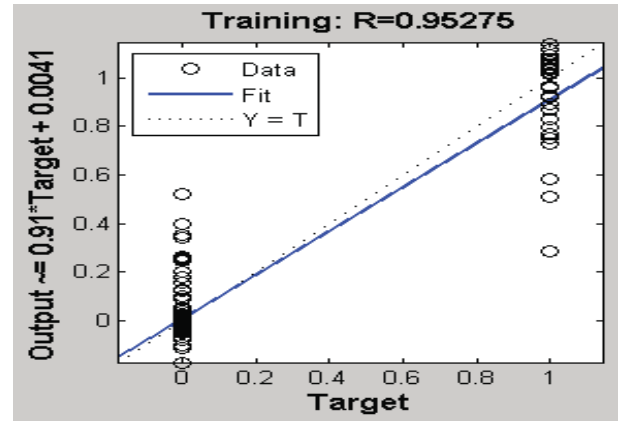


Figura 9. Gráfica de resultado de los datos de entrenamiento.

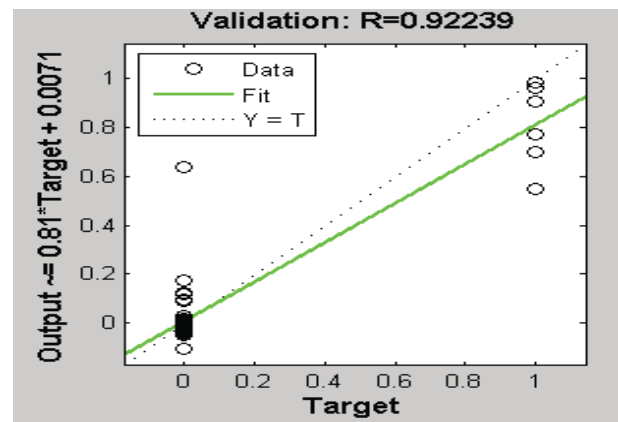


Figura 10. Gráfica de resultado de la validación de los datos.

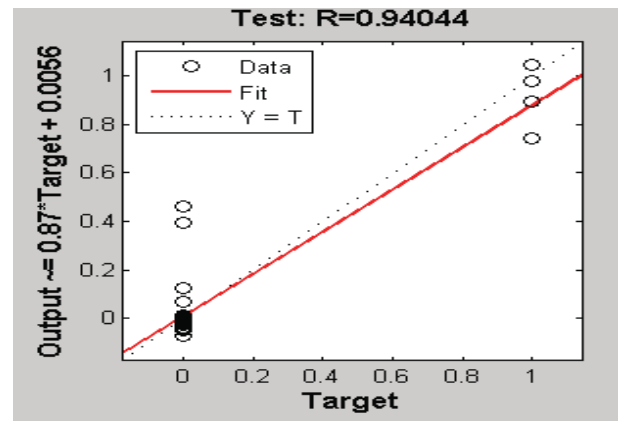


Figura 11. Gráfica de resultado de los datos de prueba.

En la última gráfica se muestra el rendimiento de la red neuronal el cual es de 0.0058613. Este rendimiento es el error cuadrático (mse) que reportó la red neuronal en Matlab después de 629 épocas. Este error es la diferencia que hay entre los datos de entrenamiento y los datos de validación.

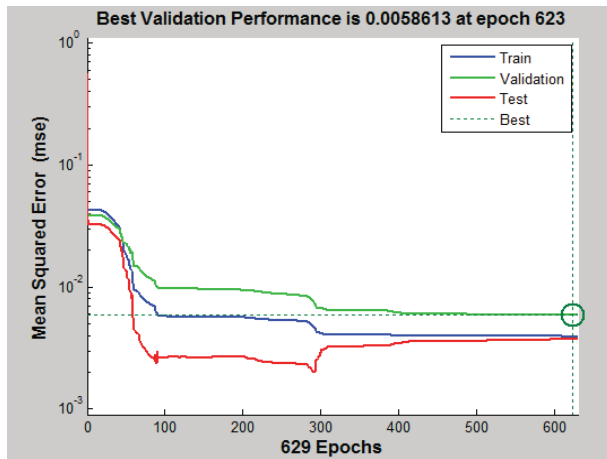


Figura 12. Gráfica de rendimiento de la red neuronal.

## V. CONCLUSIÓN Y TRABAJO FUTURO

Este trabajo presentó una solución al reconocimiento de un tipo específico de arma de fuego, con esto no se quiere decir que sea un proceso efectivo para otros tipos de armas. Dado que en el procesamiento de imágenes se realizó un algoritmo a la medida del problema que se está atacando. Así también las imágenes tomadas fueron realizadas en un ambiente semicontrolado. Es posible que en una imagen tomada al aire libre se tenga que adecuar el algoritmo y realizar cálculos extras para obtener el resultado requerido.

Este es un trabajo que forma parte de un proyecto que pretende detectar situaciones de riesgo dentro de una imagen, la cual será obtenida de un video. Pero para poder llegar a esto se tiene que realizar la detección de una persona y a su vez verificar si ésta tiene en su poder un arma de fuego, todo esto se pretende realizar en tiempo real, lo cual se logrará con ayuda de otras tecnologías como pueden ser el uso de GPU's o FPGA's, esto como trabajo futuro. Para este trabajo se pretende agregar la visión estereoscópica como un discriminante más, dado que con ésta podemos saber las dimensiones del arma y así aplicar una discriminación por tamaño. Se pretende utilizar un FPGA para realizar los cálculos que conllevan un alto costo computacional.

## REFERENCIAS

- [1] Simon Haykin. *Neural Networks and Learning Machines*. Perason, Prentice Hall, 3 edition, 2009.
- [2] ICESI. ENSI-7 resultados nacionales por entidades federativas. Technical report, Instituto ciudadano de estudios sobre la inseguridad a.c., 2010.
- [3] Michal Grega, Seweryn Lach, and Radoslaw Sieradzki. Automated Recognition of Firearms in Surveillance Video. *IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision Support (CogSIMA)*, pages 45–50, 2013.
- [4] J Carmona Suárez, I. Santillan, S. Martínez Díaz, J. L. Gómez Torres, and A. I. Barranco Gutiérrez. Procesamiento de imágenes mediante operadores de morfología matemática implementados en un fpga. In *Nuevos avances en robótica y computación CIRC 2015*, volume 1, pages 32–36, 2015.
- [5] G. J. Rodríguez, J. L. Gómez Torres, A. L. Barranco, S. Martínez, and J. Sandoval. Reconocimientos y localización visual 3d de objetos para robots manipuladores. In *CIRC 2013*, volume 1, pages 217–222, 2013.

- [6] Israel Santillan, Iván R. Terol-Villalobos, and Gilberto Herrera Ruiz. Color Morphological Image Segmentation on a New Opponent Color Space Based on the Notion of Critical Functions. *Seventh Mexican International Conference on Artificial Intelligence*, pages 213–219, 2008.
- [7] Juan H. Azuela Sossa. *Rasgos Descriptores para el Reconocimiento de Objetos*. Centro de Investigación en Computación Instituto Politécnico Nacional, 1 edition, 2006.

# Contador de colonias bacterianas en placas Petri utilizando procesamiento digital de imágenes

Mauricio Quintero Espinoza<sup>1</sup>, Jorge Manuel Santa Bárbara Cebada<sup>1</sup>, Jorge Alberto Méndez Sánchez<sup>1</sup>, Arturo Páez López<sup>1</sup>, Eliúh Cuecuecha Hernández<sup>1</sup>, Aldrin Barreto Flores<sup>1</sup>, Carlos Cabrera Maldonado<sup>2</sup>, Verónica Edith Bautista López<sup>3</sup>, Salvador E. Ayala Raggi<sup>1</sup>, Carlos Escamilla Weinmann<sup>4</sup>

Facultad de Ciencias de la Electrónica<sup>1</sup>

Facultad de Ciencias Química<sup>2</sup>

Facultad de Ciencias de la Computación<sup>3</sup>

Bioterio Claude Bernard<sup>4</sup>

Benemérita Universidad Autónoma de Puebla

Av. San Claudio y 18 Sur, Ciudad Universitaria, Puebla, Puebla, México.

[quintero.201136210@gmail.com](mailto:quintero.201136210@gmail.com)

**Abstract**— Con el fin de reducir la imprecisión y el tiempo que consume llevar a cabo conteos de colonias de bacterias en placas Petri, se propone un sistema para el recuento de colonias a través de métodos de procesamiento de imágenes con el software Matlab. Para obtener una imagen útil de las muestras se propone una estructura formada por una cámara y una plataforma con iluminación uniforme y controlada que permita capturar la imagen en las condiciones óptimas para el procesamiento. Los resultados del conteo son mostrados en una interfaz desarrollada en el mismo software. El rendimiento del sistema propuesto se verifica con recuentos manuales de un técnico con experiencia en experimentos biológicos.

**Keywords**— Colonias de bacterias, placas Petri, procesamiento digital de imagen.

## I. INTRODUCCIÓN

Actualmente el conteo bacteriano es una herramienta fundamental en investigaciones que involucran microorganismos, ya que permite cuantificar el crecimiento de bacterias en distintos medios [1]. Dichas cuantificaciones hacen evidente el comportamiento específico de los microorganismos en ambientes selectos y tiempos controlados para tomar medidas preventivas ante su reproducibilidad en caso de ser bacterias indeseables. Por ejemplo, en diversos estudios microbiológicos como propagación de virus, análisis de alimentos, de agua de bebida o de productos farmacéuticos se requiere conocer el número de microorganismos presentes en un medio particular con objeto de determinar su calidad [2]. El crecimiento de una población bacteriana puede ser entendido desde diferentes perspectivas y de acuerdo a estas se puede llegar a determinar la medida del crecimiento mediante diversas metodologías, de acuerdo con fisiólogos bacterianos, microbiólogos y expertos del área el "Recuento en placa por siembra en superficie" es uno de los métodos de conteo más sencillos y baratos que existen [3], éste consiste en la siembra de un volumen conocido de la dilución de la muestra sobre la

superficie de la placa Petri cubierto con un medio de cultivo específico haciendo que las colonias de bacterias crezcan en medio del agar, tal como se aprecia en la Fig. 1.

Esta técnica se utiliza para el recuento de bacterias tipo BMA. (Bacterias Mesofílicas Aerobias) Dicho tipo de bacterias presentan una característica de crecimiento fundamental para el desarrollo de la propuesta, esta es que el medio de cultivo es translúcido y las colonias formadas en toda la placa son de un color sólido, haciendo posible la identificación visual entre el medio y la colonia de manera clara. Este método de conteo y otros como "Recuento directo de bacterias en microscopio" o "Recuento en filtros de membrana" dependen mucho de la percepción y habilidades propias del especialista que realiza el conteo por lo que la repetibilidad de los conteos y estudios no está garantizada al cien por ciento [4]. La naturaleza del conteo de colonias bacterianas genera un proceso muy tedioso para los investigadores en el área de la microbiología, debido a que el procedimiento usual es colocar la caja Petri en un ambiente iluminado bajo una lupa que permita observar mejor la muestra. Posteriormente se procede a contar todas las bacterias observadas por el investigador y son marcadas con un plumón mediante un punto; es importante mencionar que para que una muestra pueda ser analizada, la cantidad de colonias bacterianas contenidas en la caja Petri debe estar en un rango de 25-250 bacterias [5], por lo que el tiempo invertido en el



Fig. 1. Colonias de Bacterias en placa Petri.



conteo y el error que se pueda generar por parte del investigador, pueden llegar a ser importantes.

Se debe mencionar también que existen medios electrónicos para hacerlo como el “Recuento directo con contadores electrónicos”, haciendo obligatoria la adquisición de un equipo seleccionado de entre una reducida variedad de contadores automatizados de colonias bacterianas que existen en el mercado que puedan reconocer y contar las colonias bacterianas rápidamente, pero son demasiado caros para laboratorios de universidades y para laboratorios que realizan estos análisis de manera esporádica.

Estos hechos generan la necesidad de proponer un sistema accesible que pueda contar las colonias bacterianas contenidas en un determinado cultivo [6], para ahorrar el tiempo que le toma a los investigadores realizar el análisis de manera manual y al mismo tiempo reducir el error que se pueda producir por un mal conteo al aumentar la concentración bacteriana en la muestra. El sistema que se expone en este trabajo toma la foto de la muestra con una cámara, lleva a cabo el procesamiento digital de la imagen y muestra los resultados en la interfaz desarrollada. Para llevar a cabo los conteos se utilizan diversas técnicas de procesamiento de imagen con el software Matlab, dichas imágenes son capturadas con una cámara de buena resolución y un ambiente de iluminación controlado desarrollado con una lámpara de iluminación LED y un difusor de luz de acrílico con características específicas.

## II. CONFIGURACIÓN DEL DISPOSITIVO

Los experimentos en este estudio se realizaron con cultivos de bacterias BMA (Bacterias Mesofílicas Aerobias) debido a que el agar (medio de crecimiento) de las placas Petri utilizada en este tipo de experimentos es transparente [7]. Esto es una característica útil porque podemos crear el contraste necesario entre colonias y el fondo utilizando una iluminación controlada correctamente; las muestras utilizadas para nuestras pruebas se incubaron de 24 a 48 horas por un técnico en un laboratorio especializado de acuerdo con la norma NMX-F-253-1977 para cultivo y conteo de este tipo de bacterias.

Para asegurar la precisión estadística del conteo, el número de colonias en las muestras fue limitado de 25 hasta 350 UFC (Unidades Formadoras de Colonias) porque esta cantidad de colonias se puede contar manualmente y realizar las comparaciones con el sistema que se propone.

### A. Iluminación

La iluminación es un factor muy importante en este contador de colonias porque la imagen obtenida y destinada al procesamiento digital de imagen requiere un alto contraste entre las colonias y el fondo. Una alta intensidad de luz puede producir pérdida en el número de colonias y bajas intensidades de luz pueden afectar la correcta visualización de las colonias de bacterias. Para proporcionar las características específicas requeridas para hacer un buen conteo, la placa se ilumina desde abajo por una lámpara LED. La colocación de un difusor (pieza de acrílico blanco) justo por encima de la lámpara es

importante porque garantiza una homogeneización de la luz [5], la estructura del arreglo se muestra en la Fig. 2.

Para visualizar mejor la región de interés en cada foto, el difusor se cubrió con una tapa; ésta tapa tiene un agujero centrado con un diámetro ligeramente mayor que el diámetro de la placa de Petri. Sobre esta tapa se coloca cada muestra para obtener las imágenes con la cámara. La cámara utilizada para capturar las imágenes tiene 2 millones de píxeles efectivos (1600 (H) x 1200 (V)) y se emplea esta resolución a la hora de aplicar los algoritmos de procesamiento digital de imágenes.

Todas las contribuciones de luz externa fueron bloqueadas con paredes de acrílico negro que cubren el sistema completo, permitiendo tener una imagen muy uniforme y únicamente iluminada con la lámpara led.

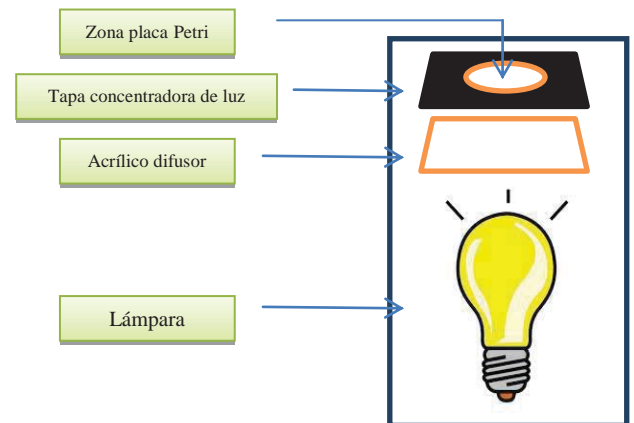


Fig. 2. Sistema de iluminación para la captura de las fotos.

## III. PROCESAMIENTO DE LA IMAGEN

Para facilitar el uso del algoritmo se hace uso de una interfaz gráfica programada en Matlab, con la finalidad de que su uso sea intuitivo y el usuario no necesite saber ningún lenguaje de programación para poder ejecutar el programa. La interfaz cuenta con un botón para su control, destacando que al seleccionarlo, se inicializa la cámara como un objeto para que Matlab tenga acceso a la cámara y pueda capturar la imagen del cultivo que se necesita [ 8]. Al finalizar el procesamiento de la imagen, el objeto creado a partir de la cámara se elimina. El botón “Iniciar” controla el momento en que el algoritmo empezara a ejecutarse y procesar la imagen. Al transcurrir escasos segundos, el resultado es arrojado de inmediato para visualizarlo en la interfaz mostrada en la Fig. 3.

### A. Captura de la imagen

Para la captura de la imagen, se hace uso de una cámara web, la cual se guarda como objeto dentro del programa para hacer uso de ella cuando se requiera. La cámara cuenta con ajustes predefinidos por el programador y estos no se modifican al correr el programa, debido a que la estructura cuenta con parámetros y condiciones controladas, para obtener condiciones semejantes al capturar cualquier imagen.

Pseudocódigo

- Inicializar la cámara como objeto
- Abrir la visualización de la cámara
- Obtener captura de la imagen de la visualización
- Guardar la imagen capturada
- Cerrar visualización de la cámara

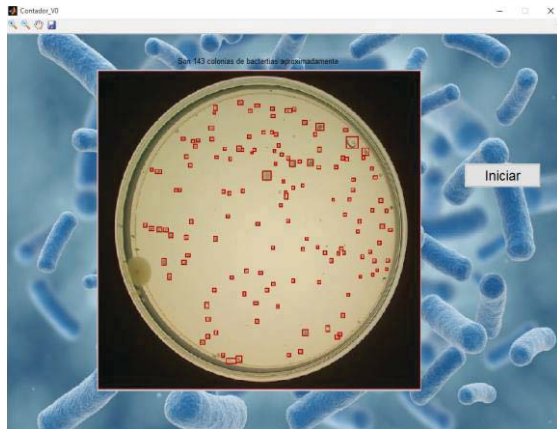


Fig. 3. Interfaz utilizada para el conteo de UFC.

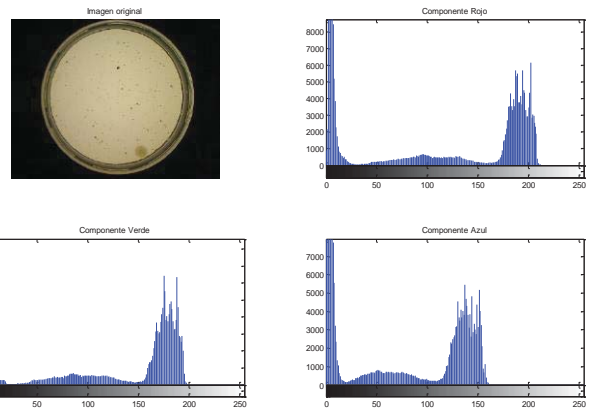


Fig. 4. Histogramas correspondientes a cada capa RGB de la imagen.

Como se desea procesar solo una imagen, se realiza una operación AND para juntar las características de las tres imágenes.

Posteriormente se erosiona la imagen binaria para tener más definidos los objetos que se desean contar, en este caso son las colonias de bacterias. La imagen binaria permite identificar con mayor facilidad los objetos presentes en la imagen, por eso se busca definir mejor los objetos debido a sus dimensiones.

Para finalizar el procesamiento de la imagen, se extraen las características necesarias tales como área y excentricidad para identificar la ubicación de los objetos y la manera de mostrar los resultados en la interfaz gráfica encerrándolos en un recuadro de color rojo.

Consideremos una matriz M de dimensiones m x n, la cual es la matriz de la imagen binaria final que se tiene del algoritmo.

$$M = \begin{bmatrix} M(1,1) & M(1,2) & \dots & M(1,n) \\ M(2,1) & M(2,2) & \dots & M(2,n) \\ \vdots & \vdots & \ddots & \vdots \\ M(m,1) & M(m,2) & \dots & M(m,n) \end{bmatrix}$$

El algoritmo se encarga de etiquetar todos los objetos encontrados con una vecindad de 4, por lo tanto, se generarán submatrices de cada uno de los objetos. Tomemos como ejemplo de una matriz generada por la identificación de un objeto N con dimensiones de p x q.

$$M = \begin{bmatrix} M(1,1) & \dots & M(1,q+2) & \dots & M(1,n) \\ \vdots & \vdots & \vdots & \dots & \vdots \\ N_{p \times q} & M(x,q+1) & M(x,q+2) & \dots & M(x,n) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ M(m,1) & \dots & M(m,q+2) & \dots & M(m,n) \end{bmatrix}$$

Donde N esta en las coordenadas de  $M(x,1)$ ,  $M(x,q)$ ,  $M(x+p,1)$  y  $M(x+p,q)$ . Estas cuatro coordenadas nos permiten conocer el área del objeto etiquetado y el centroide dentro de la matriz. Y así ubicar

B. Procesamiento de la imagen

Se ocupa el método de umbralización, donde se busca segmentar la imagen deseada. La imagen capturada es a color, por lo tanto, el algoritmo las segmenta en sus componentes RGB (Rojo, Verde y Azul). Al separar cada una de las capas en 3 imágenes diferentes, se procede a aplicarles un umbral para convertirlas a imágenes binarias.

Para obtener los umbrales adecuados en cada uno de sus componentes se procedió a obtener sus respectivos histogramas [9, 10], esto para conocer el nivel de intensidad de todos los pixeles contenidos en la matriz de cada una de las componentes, debido a que el histograma de cada componente cuenta con niveles de intensidad diferentes como se aprecia en la Fig. 4. Estos umbrales nos permitirán clasificar los elementos de la imagen en dos clases, aquellos que se encuentran por arriba de este valor serán considerados como una clase, y el resto será considerado para la segunda clase.

$$B = A > \mu \tag{1}$$

Donde la matriz B estará integrada por unos y ceros. Unos en aquellas posiciones donde los valores de los pixeles son mayores del umbral  $\mu$ , mientras que serán ceros en aquellas posiciones donde no se cumpla siguiendo la ecuación (1).

correctamente el rectángulo que encierre a cada colonia de bacterias detectada. Conociendo las coordenadas sobre la imagen es fácil pintar un recuadro, mandando un valor de color sobre las líneas que unen las 4 coordenadas encontradas.

#### Pseudocódigo

- Leer la imagen capturada
- Guardar la imagen como objeto
- Segmentar la componente roja (red) de la imagen
- Segmentar la componente verde (green) de la imagen
- Segmentar la componente azul (blue) de la imagen
- Convertir a imagen binaria la componente roja con un umbral determinado
- Convertir a imagen binaria la componente verde con un umbral determinado
- Convertir a imagen binaria la componente azul con un umbral determinado
- Realizar operación AND en las tres imágenes binarias
- Guardar la imagen resultante SUMA
- Abrir la imagen SUMA
- Definir los objetos existentes
- Realizar operación NOT en la imagen binaria SUMA
- Guardar la imagen resultante FINAL
- Abrir la imagen FINAL
- Contar el número de objetos en la imagen binaria FINAL
- Extraer características de la imagen
- Guardar características de la imagen
- Usar características para dibujar un recuadro en cada objeto localizado.
- Visualizar el resultado del conteo
- Visualizar imagen con objetos remarcados

#### IV. RESULTADOS Y DISCUSIÓN

Para evaluar el desempeño del algoritmo propuesto, imágenes de las colonias fueron contadas automáticamente aplicando la interfaz desarrollada y al mismo tiempo fueron contadas manualmente por un experto en experimentos microbiológicos. Los conteos se pueden observar en la tabla 1.

Para cuantificar el error de conteo fue necesario descartar los falsos positivos, es decir, conteos que no corresponden a una colonia de bacterias. La tonalidad de las contaminaciones es más tenue, por lo que la mayoría son eliminadas en la umbralización de las capas RGB. De acuerdo con el experto, si un cultivo de bacterias del tipo BMA posee contaminaciones, dichas formas crecen sobre la superficie del cultivo y son de

mayor tamaño que las colonias de bacterias que interesa contabilizar. Esto permite entonces discriminar las contaminaciones por el tamaño del recuadro en el que están encerradas si es que fueron detectadas como colonias de bacterias [11, 12]. De acuerdo con una serie de pruebas el área promedio del recuadro de una bacteria es 20 X 20 pixeles .

Por lo tanto el criterio para descartar falsos positivos, midiendo el área de los recuadros obtenidos al marcar las colonias detectadas es el siguiente:

$$\text{Area del recuadro} > 500 \text{ px} = \text{Falso positivo.}$$

La precisión del sistema propuesto en base a 6 placas contadas es buena ya que en todos los casos obtuvimos un conteo mayor al de las cuentas manuales realizadas por personal experto, en las cuales el sistema detectó colonias muy pequeñas. En un conteo manual las colonias ubicadas en la periferia son despreciadas, ya que por la ubicación de las colonias y la estructura física de las placas Petri se dificulta la visión de las mismas. Tal como se aprecia en la Fig.5.

El proceso de conteo arrojó los siguientes resultados.

TABLA 1. COMPARACIÓN ENTRE CONTEOS MANUALES Y CONTEOS DEL ALGORITMO PROPUESTO. ESTOS RESULTADOS MUESTRAN 6 MUESTRAS ANALIZADAS

No. de Muestra	Conteo Manual	Conteo del algoritmo propuesto
1	132	139
2	168	180
3	248	265
4	94	107
5	82	92
6	48	56

Tomando en cuenta los resultados obtenidos de la tabla anterior, el sistema propuesto presenta una precisión buena ya que en el índice de conteo de colonias de bacterias en placas con concentraciones de 50 a 300 UFC se mantiene en un rango de 5 a 20 colonias más comparadas con las cuentas manuales. Este rango de acuerdo con un técnico experto en experimentos biológicos es muy bueno.

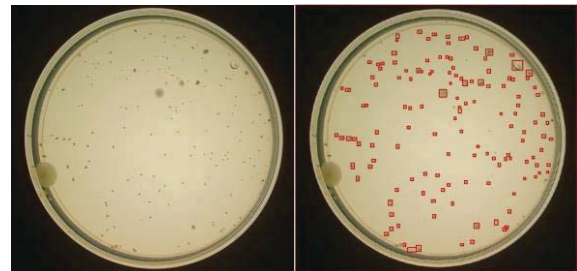


Fig. 5. Conteo del algoritmo propuesto sin restar los falsos positivos.

#### V. CONCLUSIONES Y TRABAJO FUTURO

En este trabajo se presentó un sistema basado en procesamiento digital de imágenes capaz de contar colonias de bacterias en

placas Petri del tipo BMA (Bacterias Mesofílicas Aerobias) con un error aceptable de acuerdo con el técnico experto en experimentos microbiológicos. Contaminaciones en los cultivos de bacterias pueden dificultar el procesamiento de la imagen, por lo que deben ser evitadas durante el procedimiento de siembra. Del mismo modo el conteo puede variar si hay imperfecciones en las placas Petri como rayones o manchas siendo preferible que las placas no estén maltratadas. El desempeño de este sistema es satisfactorio solo para el tipo de bacteria propuesto, ya que otro tipo de cultivos, principalmente los de medio de cultivo de color, no son reconocidos correctamente por el algoritmo de procesamiento. Desde este punto de vista, como trabajo futuro se pretende ampliar la capacidad de conteo del dispositivo desarrollando diferentes tipos de algoritmos para las necesidades presentes en los laboratorios, así como ampliar la capacidad de conteo de placas de muestras a más de 1.

El desempeño de este sistema es satisfactorio solo para el tipo de bacteria propuesto, ya que otro tipo de cultivos, principalmente los de medio de cultivo de color, no son reconocidos correctamente por el algoritmo de procesamiento. Desde este punto de vista, como trabajo futuro se pretende ampliar la capacidad de conteo del dispositivo desarrollando diferentes tipos de algoritmos para las necesidades presentes en los laboratorios, así como ampliar la capacidad de conteo de placas de muestras a más de 1.

#### REFERENCIAS

- [1] Hüseyin Ates, Ömer Nezh Gerek, "An Image-Processing Based Automated Bacteria Colony Counter", IEEE , 2009.
- [2] M. Niyazi, I. Niyazi, and C. Belka, "Counting colonies of clonogenic assays by using densitometric software," Radiation Oncology, vol. 2, pp. 4, 2007.
- [3] J. Marotz, C. Lübbert, W. Eisenbeib, "Effective object recognition for automated counting of colonies in Petri dishes (automated colony counting)," Computer Methods and Programs in Biomedicine, vol. 66, pp. 183-198, 2000.
- [4] Chengcui Zhang, Wei-Bang Cheng, Wen-Lin Liu, Chi-Bang Chen "An Automated bacterial colony counting system," IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing , 2008.
- [5] Pei-Ju Chiang, Min-Jen Tseng, Zong-Sian He, Chia-Hsun Li, "Automated counting of bacterial colonies by image analysis," Journal of Microbiological Methods , vol. 108, pp. 74-82, 2015.
- [6] C. W. Chang, Y. H. Hwang, S. A. Grinshpun, J. M. Macher, and K. Willeke, "Evaluation of Counting Error Due to Colony Masking in Bioaerosol Sampling," Applied and Environmental Microbiology, vol. 60, pp. 3732-3738, 1994.
- [7] X. Liu, S. Wang, L. Sendi, and M. J. Caul\_eld, "High- throughput imaging of bacterial colonies grown on filter plates with application to serum bactericidal assays," Journal of Immunological Methods, vol. 292, pp. 187-193, 2004.
- [8] Rafael C. Gonzalez, Richard E. Woods and Steven L. Eddins. Digital Image Processing Using MATLAB. United State: Prentice Hall. 2004.
- [9] Oscar Möller, Javier W. Signorelli, Mario A. Storti (Eds.), "Métodos de umbralización de imágenes digitales basados en entropía de Shannon y otros", Mecánica Computacional Vol XXX, págs. 2785-2805 (artículo completo) Rosario, Argentina, 1-4 Noviembre 2011.
- [10] MATLAB. Analyzing the texture of an image. Image Processing Toolbox User's Guide. 2010. 417-426.
- [11] M. Putman, R. Burton, and M. H. Nahm, "Simplified method to automatically count bacterial colony forming unit," J Immunol Methods, vol. 302, pp. 99-102, 2005.
- [12] MathWorks. regionprops – Pixel Values and Statistics. *Image Processing Toolbox*

# Sistema mecatrónico para la identificación de dispositivos en tarjetas PCB usando visión por computadora

Juan Carlos Campeche Valencia<sup>1</sup>, Aldrin Barreto Flores<sup>1</sup>, Verónica Edith Bautista López<sup>2</sup>, Carlos García Lucero<sup>1</sup>, Salvador E. Ayala Raggi<sup>1</sup>

Facultad de Ciencias de la Electrónica<sup>1</sup>

Facultad de Ciencias de la Computación<sup>2</sup>

Benemérita Universidad Autónoma de Puebla

Av. San Claudio y 18 Sur, Ciudad Universitaria, Puebla, Puebla, México

Campeche.200928008mail.com

**Abstract—** The automated optical inspection (AOI) in the electronic manufacturing industry is a critical process to reduce manufacturing cost and assembly issues. This paper describes a mechatronic system that is able to identify if an electronic device or component is on a printed circuit board (PCB) using machine vision techniques, also a conveyor belt is used to automate the inspection process. The system mainly include a illumination system, an image processing module, algorithms of machine vision based on OpenCV and implemented in a Raspberry Pi, an user interface made in QT Creator and a microcontroller that is employed to automate the system. This system is a low cost alternative in the non-contact and non-destructive inspection systems.

**Keywords—** Automated optical inspection; machine vision; Raspberry PI.

## I. INTRODUCCIÓN

En el proceso de manufactura de un circuito impreso existen principalmente 2 tipos de inspecciones para encontrar fallas: las de contacto eléctrico y sin contacto eléctrico. La inspección que no implica contacto eléctrico evalúa la ausencia, la mala colocación de dispositivos, exceso o falta de soldadura y el tamaño de canales conductores [1]. Esta inspección se puede realizar por personas pero el proceso es muy lento y es muy susceptible a errores humanos debido al cansancio y a las pequeñas dimensiones de los dispositivos, es por esto que se han implementado sistemas de visión por computadora para realizar una inspección que garantice la presencia de los dispositivos. La inspección de tarjetas PCB es un campo en crecimiento, ya que la industria cuenta con alrededor de 2800 empresas y continúan creciendo [2], lo que ocasiona que existan investigaciones como en [3] donde se verifica la posición de resistencias superficiales o [4] donde se usa aprendizaje basado en programación genética para

encontrar los posibles defectos. Lamentablemente estos sistemas son muy exclusivos debido a su alto costo por lo que solo pueden ser adquiridos por empresas con grandes niveles de producción, por esta razón se propone el desarrollo de un sistema basado en software Open Source, como son las bibliotecas de visión por computadora de OpenCV y el entorno de desarrollo del lenguaje C++ como QT Creator, además de que el sistema estará soportado en una Raspberry Pi lo que lo convierte en un sistema pequeño y económico.

## II. METODOLOGÍA

Para realizar el sistema de inspección automática óptica se desarrolló un sistema de visión por computadora compuesto por una cámara digital (RaspiCam), una Raspberry Pi, un sistema de iluminación, un servomotor, un sensor infrarrojo, un microcontrolador Arduino Leonardo y una banda transportadora. La tarjeta PCB a ser evaluada es una tarjeta de desarrollo para microcontroladores, la cual tiene una gran diversidad de dispositivos que pueden ser identificados. Estas tarjetas PCB se colocan sobre la banda transportadora, posteriormente un sensor infrarrojo detecta la presencia de una tarjeta y un servomotor posiciona la tarjeta para iniciar la toma de imágenes. Se capturan las imágenes con un fondo blanco para elevar el contraste, un entorno cerrado para evitar la entrada de luz ambiental y tener una iluminación controlada. La fuente de luz son LEDs blancos difusos y se evita la iluminación frontal porque las tarjetas PCB son superficies altamente reflejantes. La primera parte del algoritmo de inspección consiste en hacer un preprocesamiento de la imagen para eliminar zonas que no son de interés y resaltar componentes importantes. En la Figura 1 se muestra el diagrama de flujo del funcionamiento general del sistema de inspección.

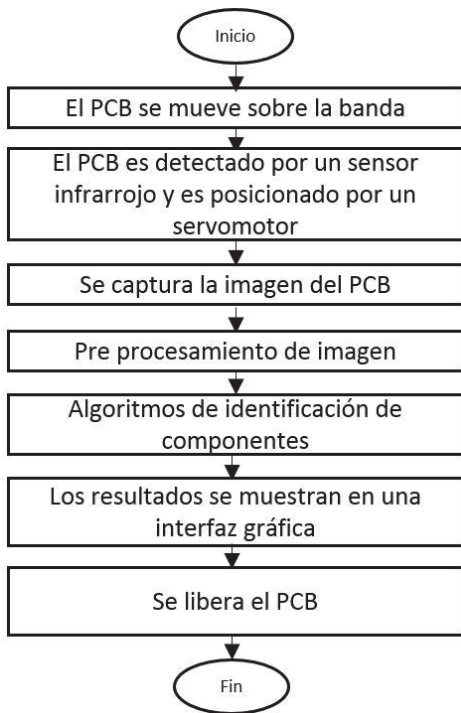


Fig. 1. Diagrama de flujo del funcionamiento general del sistema

### A. Preprocesamiento de imagen

El preprocesamiento de la imagen nos permite eliminar los elementos que dificultan la identificación, así como resaltar los elementos importantes como son transistores, resistencias y circuitos integrados [5]. Esta etapa se realiza en cuatro pasos: primero se extrae de la imagen tomada (tamaño 2592x1944) únicamente la tarjeta PCB con el objetivo de tener una imagen pequeña de 1250x740 píxeles para disminuir tiempos de procesamiento. Después se aplica un filtro Gaussiano [6] con un núcleo de 5x5 para suavizar la imagen pero sin diluir los bordes. En el paso siguiente se realiza una conversión de modelo de color RGB (Modelo por defecto) a un modelo de color HSV o también conocido como HSL [7], el cual es el modelo que permite la segmentación basada en color, ya que representa la información del color a través de 3 componentes: la tonalidad (Hue), la saturación (Saturation) y valor (Value), lo que permite eliminar el color azul de fondo del PCB y así elevar el contraste de los componentes. En el siguiente paso se hace un umbralado superior e inferior (Eq. 1) en cada componente del modelo HSV para remover el color de la placa y dejar únicamente una imagen binaria, la cual servirá de máscara para resaltar los dispositivos como se muestra en la Figura 2. Los valores de umbralado que permiten remover el color de fondo son  $H_{inf} = 31$  y  $H_{sup} = 122$  para la tonalidad,  $I_{inf} = 97$  y  $I_{sup} = 255$  para la saturación y  $H_{inf} = 0$  y  $H_{sup} = 200$  para la iluminación. El último paso es aplicar la máscara obtenida a la imagen RGB de la tarjeta PCB como se muestra en la Figura 2.

$$Um(I_{HSV}(x, y)) = H_{inf} \leq I_H \leq H_{sup} \cap S_{inf} \leq I_S \leq S_{sup} \cap V_{inf} \leq I_V \leq V_{sup} \quad (1)$$

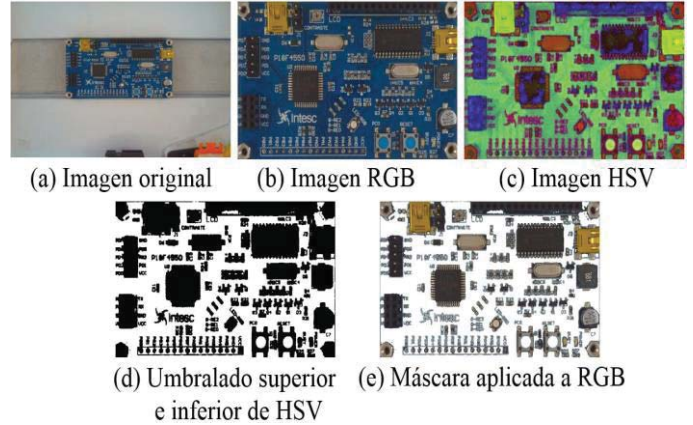


Fig. 2. Etapas de preprocesamiento de imagen

### B. Identificación de dispositivos

El proceso para identificar la presencia de dispositivos se basa en la técnica de comparación de imágenes (Template Matching) [6], donde una imagen patrón (Template) se mueve por todos los píxeles de una imagen para encontrar la máxima coincidencia como se observa en la Fig. 3. Existen diferentes métodos para realizar esta técnica pero en este trabajo se utilizó el método Coeficiente de Correlación Normalizado (NCC) o Coeficiente de correlación lineal de Pearson [8] definido por la Eq. 2, donde  $C_L(r, s)$  es la NCC,  $R(i, j)$  es la imagen patrón,  $I(i, j)$  es la imagen analizada,  $r$  y  $s$  son el desplazamiento de  $R(i, j)$  en  $I(i, j)$ ,  $\bar{I}(r, s)$  y  $\bar{R}$  son las diferencias locales en cada operación. Se eligió este método porque proporciona la ubicación  $(x, y)$  de las coincidencias de la imagen patrón, lo cual es indispensable para hacer la identificación automática de los diferentes dispositivos. Además este método puede hacer la localización aun cuando existan variaciones en la inclinación de las imágenes de hasta 10 grados, cuando se excede esta inclinación este método no hace detecciones correctas.

$$C_L(r, s) = \frac{\sum_{(i,j) \in R} (I(r+i, s+j) - \bar{I}(r, s)) * (R(i, j) - \bar{R})}{\left[ \sum_{(i,j) \in R} (I(r+i, s+j) - \bar{I}(r, s))^2 \right]^{\frac{1}{2}} \left[ \sum_{(i,j) \in R} (R(i, j) - \bar{R})^2 \right]^{\frac{1}{2}}} \quad (2)$$

Donde  $\bar{I}(r, s)$  y  $\bar{R}$  se definen como:

$$\bar{I}(r, s) = \frac{1}{N} \sum_{(i,j) \in R} I(r+i, s+j) \quad \bar{R} = \frac{1}{N} \sum_{(i,j) \in R} R(i, j) \quad (3)$$

El método NCC obtiene valores en el intervalo  $[-1, +1]$ , donde el valor +1 indica una coincidencia perfecta y a menor coincidencia el valor disminuye hasta -1, como se muestra en la Figura 3. Este método es una variante mejorada del

Coeficiente de Correlación, ya que compensa cambios de brillo y contraste en la imagen

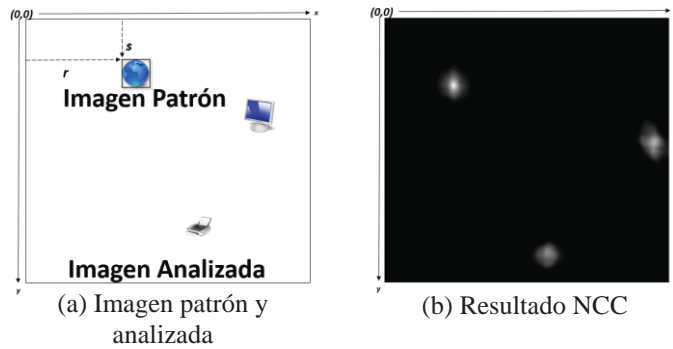


Fig. 3. Método Template Matching

La aplicación del método NCC se realiza dividiendo la imagen en zonas de interés para reducir el procesamiento, ya que no se pretende buscar un transistor donde se conoce que no hay transistores. El siguiente paso es hacer un umbralado a cero [9] como se indica en la Eq. 4 al resultado de NCC con el objetivo de eliminar los dispositivos que no son lo suficientemente parecidos a la imagen patrón y los que superen el umbralado son los que tienen un gran parentesco a la imagen patrón (Figura 4).

$$Um(Im(x,y)) = \begin{cases} Im(x,y) & \text{si } Im(x,y) \geq Umbral \\ 0 & \text{si } Im(x,y) < Umbral \end{cases} \quad (4)$$

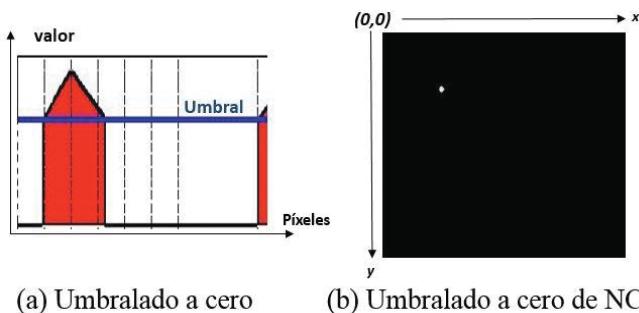


Fig. 4. Segmentación por umbral

Por ultimo comparamos la posición (x,y) de todos los candidatos obtenidos anteriormente con una tabla de referencia generada de manera manual, la cual contiene todas las posiciones (x,y) correctas de cada dispositivo considerando un círculo de tolerancia de 5 pixeles, ya que también existe una tolerancia de posición para cada dispositivo colocado.

### C. Sistema mecatrónico desarrollado

El desarrollo del sistema mecatrónico se basa en un microcontrolador Arduino Leonardo, el cual recibe la señal de un sensor infrarrojo SHARP 2Y0A21 cuando se detecta la presencia de una tarjeta PCB, al recibir la señal el microcontrolador le indica a un servomotor que mueva un mecanismo, el cual mantiene la posición de la tarjeta PCB

mientras se hace la toma de imagen y después la identificación de dispositivos. Después el microcontrolador por medio de comunicación serial le indica a la Raspberry Pi modelo B que tome la foto y que inicie el algoritmo de identificación. Cuando se termina el algoritmo se muestran los resultados en una interfaz gráfica desarrollada en Qt Creator y se le indica al microcontrolador que libere a la tarjeta PCB para esperar la siguiente tarjeta a analizar.

El sistema se propuso como se muestra en la Figura 5 porque se necesita tener una imagen frontal de la tarjeta PCB y una iluminación controlada, la cual se obtiene cerrando un área determinada para evitar la entrada de luz ambiental y poniendo un sistema de iluminación. El sistema de iluminación está compuesto de dos placas LED, las cuales generan una luz blanca difusa para evitar reflejos en la imagen.

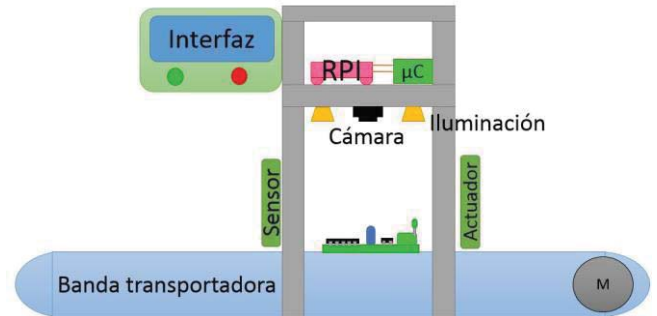


Fig. 5. Diagrama del Sistema mecatrónico

### III. DISCUSIÓN DE RESULTADOS

Las pruebas se realizaron con el sistema que se muestra en la Figura 6. Debido a que la identificación de dispositivos se basa en la técnica de comparación de imágenes se usaron imágenes patrón como los que se muestran en la Figura 7.

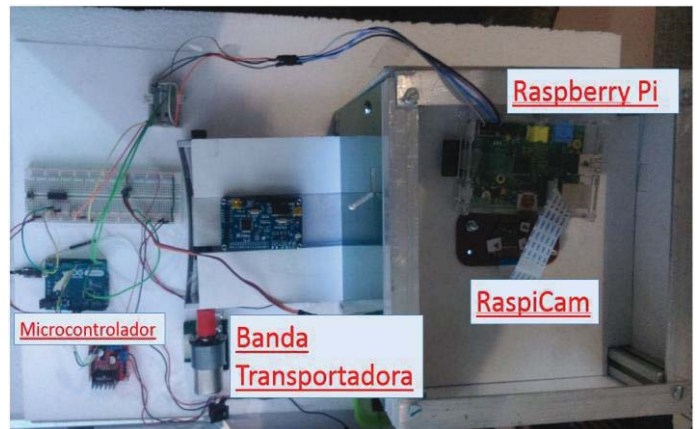
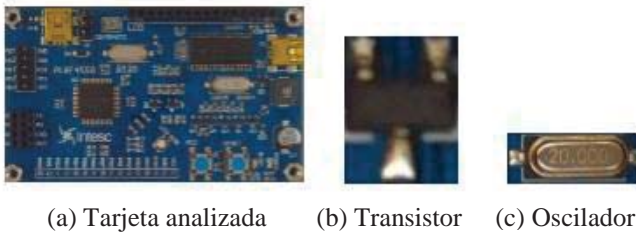


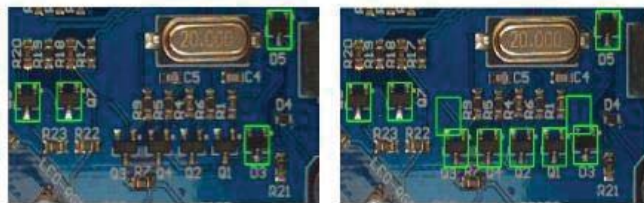
Fig. 6 Sistema mecatrónico para la identificación de dispositivos



(a) Tarjeta analizada (b) Transistor (c) Oscilador

Fig. 7. Imagen analizada y patrón

La técnica de comparación de imágenes permite encontrar un patrón dentro de otra imagen, sin embargo esta técnica es muy susceptible a cambios de forma, iluminación y tamaño en la imagen analizada. Para solucionar este problema se mantuvo una iluminación constante, se implementó una banda transportadora para mantener la tarjeta en posición horizontal y un servomotor mantiene la posición durante el análisis. Aplicando únicamente estas restricciones y haciendo umbralado de 0.7 después del método NCC se obtuvo un porcentaje de efectividad en la identificación de dispositivos del 50% (Figura 8), modificando el umbralado a 0.5 se encuentran todos los dispositivos pero también tiene identificaciones erróneas. Estos porcentajes se deben a que existen pequeñas diferencias de color alrededor del dispositivo lo que ocasiona que la comparación de imágenes tenga resultados indeseados.



(a) Umbralado = 0.7 (b) Umbralado = 0.5

Fig. 8. Identificación de dispositivos sin preprocesamiento de imagen

Para mejorar los resultados se implementó la etapa de preprocesamiento descrita anteriormente, además la imagen patrón fue modificada para eliminar las variaciones de colores del fondo como se muestra en la Figura 10. Con estas modificaciones y un umbralado de 0.7 después del método NCC se obtuvo una efectividad del 100% en la identificación, lo que nos demuestra que se pueden contrarrestar algunas deficiencias del método de comparación de imágenes con un preprocesamiento adecuado.

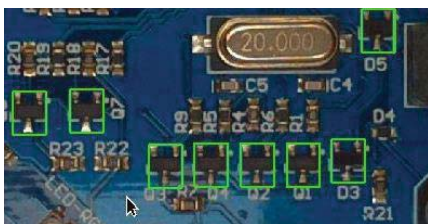
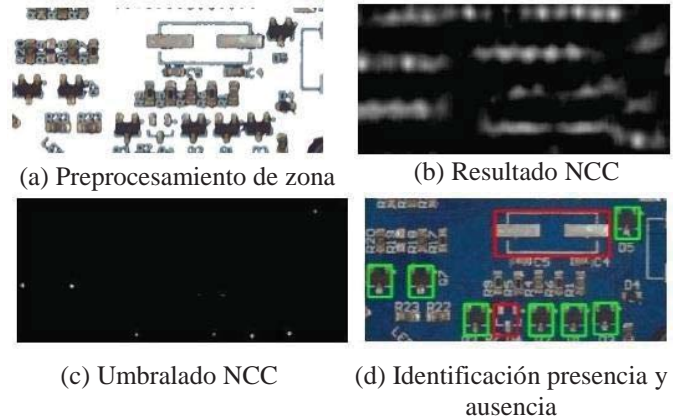


Fig. 9. Con preprocesamiento y umbralado de 0.7

La siguiente etapa fue evaluar tarjetas PCB con diferentes dispositivos faltantes, para hacer la identificación de diferentes

dispositivos se necesita realizar el método NCC el mismo número de veces que dispositivos, así como realizar diferentes umbralados para cada dispositivo. A continuación se muestra el resultado de buscar dos diferentes dispositivos.



(a) Preprocesamiento de zona (b) Resultado NCC  
(c) Umbralado NCC (d) Identificación presencia y ausencia

Fig. 10. Identificación de diferentes dispositivos

El sistema mecatrónico cuenta con una interfaz gráfica desarrollada en QT Creator, desde la cual se controla todo el sistema y muestra los resultados de la inspección, como se muestra en la Figura 11.

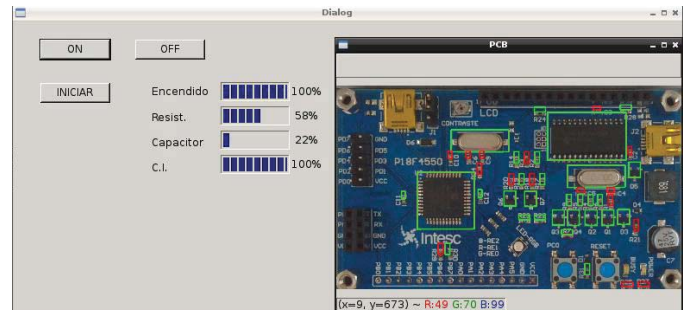


Fig. 11. Interfaz gráfica del sistema

En la interfaz gráfica se presenta con color verde si un dispositivo está presente y si está es su posición correcta, con color rojo se muestra la posición de un dispositivo ausente. En las barras de progreso se muestra el porcentaje de cada dispositivo encontrado. Los resultados obtenidos después de realizar pruebas a 5 diferentes tarjetas con diferentes errores se muestran en la Tabla 1. Podemos apreciar que para dispositivos con formas y colores muy definidos como son los transistores, osciladores, botones y circuitos integrados se obtiene un porcentaje de efectividad del 100%, es decir, el sistema es capaz de identificar perfectamente si el dispositivo está o no. Sin embargo, para dispositivos pequeños como resistencias y capacitores se obtiene un porcentaje de reconocimiento del 80%, esto se debe a que los dispositivos tienen un tamaño muy pequeño (15 x 25), además su composición de colores puede estar contenida dentro de un transistor o ser muy parecida a la de un capacitor, por lo cual se debe mejorar el preprocesamiento o usar una técnica de identificación únicamente para estos elementos. Esto ocasiona que se detecten falsos positivos cuando una resistencia o un



capacitor están ausentes debido a que existe muy poca diferencia cuando el dispositivo está o no.

Dispositivo	Porcentaje de identificación	Dispositivo	Porcentaje de identificación
Transistores	100%	Diodos	70%
Osciladores	100%	Circuitos	100%
Resistencias	80%	Integrados	
Capacitores	80%		

Tabla 1 Resumen de resultados

#### IV. CONCLUSIONES

La producción de sistemas electrónicos requiere una inspección confiable y rápida para evitar errores que ocasionen devoluciones y retrabajos. Sin embargo los sistemas de visión para inspección automática son muy costosos lo que los hace inaccesibles para muchas empresas medianas o nuevas. La inspección automática óptica basada en sistemas embebidos y software Open Source presenta una gran alternativa para realizar inspecciones confiables a un bajo costo, ya que los componentes de este sistema no exceden los 120 USD. El sistema también puede ser utilizado y reconfigurado para verificar al momento de entrega los sub ensambles que se adquieran de proveedores. Este trabajo posee las características de ser una plataforma para diversificar los objetos a analizar y además de servir como una herramienta de aprendizaje e investigación. El trabajo a futuro incluye mejorar el algoritmo para identificar resistencias, además se pretende abarcar otros modelos de tarjetas para

hacer más robusto el algoritmo y la verificación de las dimensiones de las pistas conductoras de la tarjeta PCB.

#### REFERENCIAS

- [1] M. Moganti, F Ercal, C. H. Dagli, and S. Tsunekawa.: Automatic pcb inspection algorithms: a survey. *Computer vision and image understanding*, vol 63, no. 2, Pag, 287-313, 1996.
- [2] F. Guo and S.-a. Guan.: Research of the machine vision based pcb defect inspection system. Pag, 472-475, 2011.
- [3] H. Wu, G. Feng, H Li, and X. Feng, Automated visual inspection of surface mounted chip components.: *Mechatronics and Automation (ICMA)*. 2010 International Conference on. Pag, 1789-1794, 2010.
- [4] F. Xie, A. Uitdenbogerd, and A. Song.: Detecting pcb component placement defect by genetic programming. *IEEE Congress on Evolutionary Computation*, Pag, 1138-1145, 2013.
- [5] M. Kumar, M. Kumar, and G. Kumar.: PCB image Enhancement Using Machine Vision For Effective Defect Detection. *International Journal of Advanced Engineering Research and Science (IJAERS)*. Pag, 50-53, 2014.
- [6] Cuevas, E., Zaldivar, D., & Pérez, M. (2010). Procesamiento digital de imágenes con MATLAB y Simulink. Alfaomega Ra-Ma (México).
- [7] Szeliski, R. (2010). *Computer vision: algorithms and applications*. Springer Science & Business Media.
- [8] D. Koniar, L. Hargas, A. Simonova, M. Hrianka, and Z. Loncova.: Virtual Instrumentation for Visual Inspection in Mechatronic Applications. *Modelling of Mechanical and Mechatronic Systems MMaMS*, Pag, 227-234, 2014
- [9] Bradski, G., and Kaehler. (2008) *Learning OpenCV: Computer vision with the openCV library*. O'Reilly Media, Inc

# Implementación de un filtro de correlación no lineal en una GPU

Toral Manriquez, Hugo  
Maestría en Sistemas  
Computacionales  
ITLP  
La Paz, B.C.S., México  
hugotm03@gmail.com

Santillán Méndez, Israel  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S., México  
santilis@gmail.com

Martínez Díaz, Saúl  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S., México  
smdiaz06@hotmail.com

Castro Liera, Iliana  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S., México  
Iliana.castro@gmail.com

**Resumen-** En este artículo se presenta una técnica para el reconocimiento de objetos que utiliza un filtro de correlación no lineal implementado en una GPU (unidad de procesamiento gráfico). Esta técnica ha demostrado ser muy eficaz en la detección de objetos dentro de imágenes ruidosas, básicamente se realiza un barrido del objeto a buscar dentro de la imagen pixel por pixel, lo que genera un costo computacional elevado, debido a esto se decidió implementar en esta tecnología paralela auxiliados de CUDA C. los resultados nos muestran una reducción en el tiempo de ejecución considerable.

**Palabras clave**—Reconocimiento de objetos; correlación no lineal; Procesamiento Paralelo; GPU; CUDA;

## I.INTRODUCCIÓN

El ser humano cuenta con una gran capacidad visual para realizar la detección de objetos sin embargo no siempre es posible contar con uno ya sea porque se trate de un lugar inaccesible, se encuentre en un ambiente con poca visibilidad o se trate de un lugar de alto riesgo, estas fueron algunas de las razones que dieron lugar a la creación de áreas de investigación que intentan resolver el problema de manera automática [1].

Dentro del área de procesamiento de imágenes y la visión por computadora existen diversos métodos para la detección de objetos sin embargo cuentan con factores que limitan su eficiencia como el ruido, la iluminación y las distorsiones geométricas, motivados por solventar estos inconvenientes se comenzó a trabajar con filtros de correlación no lineal ya que demostraron ser robustos ante el ruido.

La correlación no lineal nos indica que tan similares o diferentes son nuestra ventana objetivo y el objeto analizado en nuestra escena de prueba utilizando medidas que no cumplen con el principio de superposición, en particular la medida aquí utilizada trabaja en el orden de  $n^2 \times m^2$  dando como resultado una técnica costosa para un procesador secuencial[2].

Afortunadamente la computación ha ido evolucionando y ya existen herramientas con procesamiento paralelo ejemplo de ello son las tarjetas gráficas o GPU por sus siglas en inglés Graphics Processors Unit, uno de sus principales fabricantes es la empresa NVIDIA quien en sus primeras versiones de desarrollo estaban enfocadas principalmente en el área de los video juegos, posteriormente fueron utilizadas para un propósito general, el propietario de estas tarjetas otorgó la

facilidad de que se utilizaran de esta manera desarrollando CUDA la cual es una plataforma de computación paralela que permite utilizar la GPU mediante CUDA C un lenguaje basado en c++ [3,4].

## II.METODOLOGIA

La técnica que se está utilizando consta de una ventana en la cual se encuentra el objeto a buscar, esta ventana se desplaza sobre cada pixel en la escena de prueba de la siguiente manera: primero, la ventana deslizante se centra en cada pixel de la escena de prueba. Después se extrae de la escena de prueba la vecindad de píxeles que rodean al píxel central y se procede con la comparación punto a punto representada por (1)[2].

$$c(m, n) = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} \text{MIN}(g(p, q), h(m + p, n + q)) \quad (1)$$

Donde  $c$  representa nuestro factor de correlación,  $g$  es nuestra ventana deslizante y  $h$  nuestra escena de prueba.

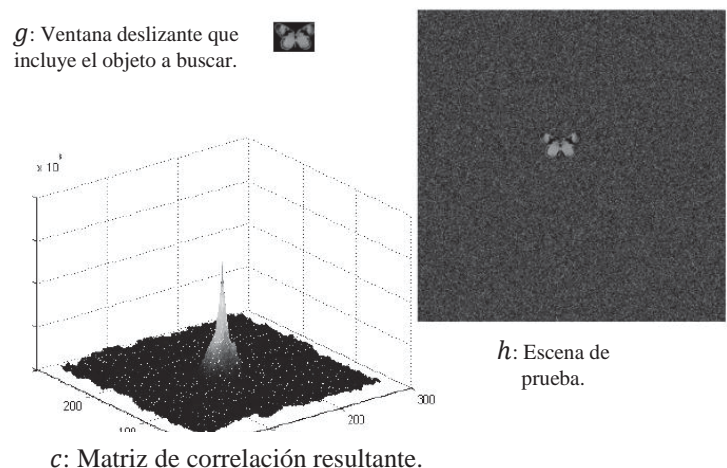


Fig.1. Componentes del filtro utilizado.

Visto lo anterior se puede apreciar que la complejidad algorítmica del filtro está definida por el tamaño de la ventana y el tamaño de la escena de prueba, por ejemplo supongamos que tenemos una ventana de tamaño  $m$  y una escena de tamaño  $n$ , la complejidad resultante será  $m \times m$  cada pixel de la ventana por  $n \times n$  cada pixel de la escena es decir  $m^2 n^2$ . Sabemos que para cada pixel se realiza exactamente el mismo tratamiento por lo tanto podemos observar que es un problema que se puede paralelizar.

### 2.1 IMPLEMENTACIÓN EN MATLAB

Para fines de comparación y veracidad en los datos obtenidos por la versión paralela del filtro es necesario realizar una implementación secuencial del mismo, para ello nos auxiliamos de una herramienta denominada MATLAB la cual es una abreviación de Matrix Laboratory o laboratorio de matrices, es un entorno de desarrollo con un lenguaje de alto nivel que cuenta con múltiples herramientas, entre ellas destaca la de gráficos con la que es posible importar y exportar imágenes así como graficar matrices[5].

Las actividades en la implementación en MATLAB son las siguientes:

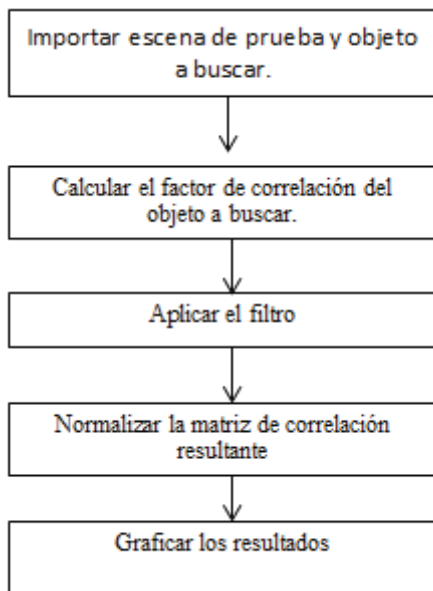


Fig. 2. Actividades realizadas en Matlab.

### 2.2 IMPLEMENTACION EN CUDA

Al ser un problema con posibilidad de paralelizar es conveniente acudir a una tecnológica con este tipo de procesamiento, adicionalmente de una GPU existen otras herramientas como lo es MPI un modelo para clúster de computadoras que no comparten memoria, Open MP para sistemas multiprocesador de memoria compartida, sin embargo al existir CUDA nos facilita el manejo de la tarjeta gráfica llevándola a un ambiente con lenguaje de alto nivel.

Como se aprecia en la fig. 3, la arquitectura CUDA se compone de hilos, bloques y una malla.

Hilos: Es donde se ejecutan las funciones, pueden realizar procesamiento simultáneamente y cuentan con memoria local para almacenar los datos.

Bloques: un conjunto de hilos que se ejecutan en un multiprocesador. Cada bloque cuenta con una zona de memoria denominada memoria compartida, accesible para todos sus hilos.

Malla: es un conjunto de bloques, solo existe una para cada GPU y cuenta con memoria global accesible para todos los bloques que lo componen[6].

Host: se le denomina host a todo lo que ocurre del lado de la CPU.

Device: recibe este nombre lo perteneciente a la GPU.

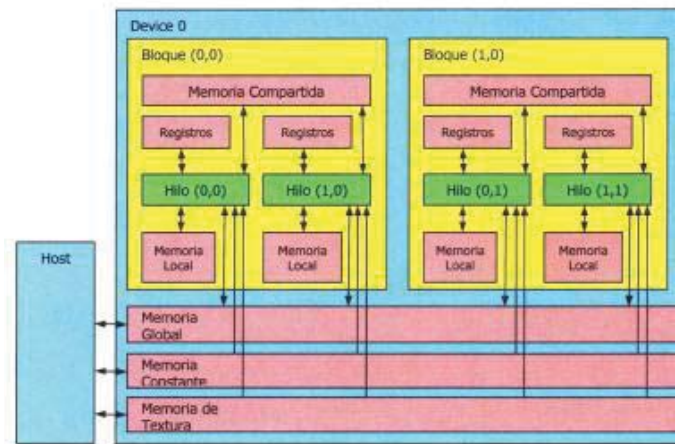


Fig. 3. Arquitectura CUDA

La tarjeta utilizada en este trabajo es una NVIDIA GeForce GTX Titan X con soporte para utilizar CUDA, para trabajar con imágenes de resolución 512 x 512, 640 x 640, 800 x 800 pixeles y ventanas de 32 x 32, 64 x 64, 128 x 128 pixeles, se optó por optimizar el uso de los hilos de procesamiento acomodando la imagen en 128 bloques de 2048 hilos cada uno, de esta manera cada hilo se encarga de realizar el cálculo del factor de correlación de su pixel correspondiente teniendo un avance de 2048 pixeles a la vez y no de 1 como lo sería en un procesador secuencial, el funcionamiento del filtro en la GPU es como se muestra en la figura 4.

La CPU utilizada es una Intel de 4ta generación i5-4440 que cuenta con 4 núcleos la frecuencia básica es de 2.8 GHz y hasta un máximo de 3.3 GHz [7].

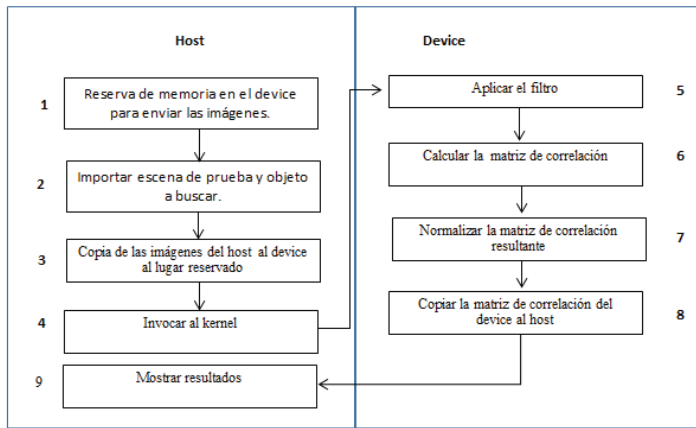
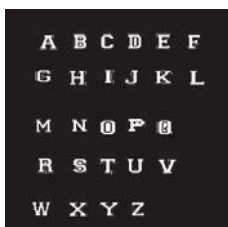


Fig. 4. Actividades realizadas en CUDA

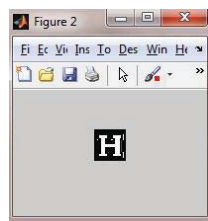
1. Se realiza la reserva de la memoria de las variables a utilizar del lado del device como lo son: la imagen de escena de prueba, la imagen con el objeto a buscar y la matriz de correlación.
2. Para la lectura de las imágenes se está utilizando la función `imread ()` de `openCV`.
3. Ya que la memoria en el device se encuentra reservada la copia de las imágenes se realiza mediante la instrucción “`cudaMemcpyToSymbol`”.
4. La invocación de un kernel se hace de la siguiente manera `filtro<<<bloques, hilos>>>()`;
5. En el lado del device cada hilo le aplica el filtro a un pixel.
6. Se calcula el factor de correlación correspondiente a cada pixel.
7. El factor se normaliza al mismo tiempo que se guarda en la variable global.
8. Se copia la matriz de regreso al host usando la instrucción “`cudaMemcpyFromSymbol`”
9. Se muestra en consola el tiempo de ejecución y los valores de la matriz de correlación.

III. RESULTADOS

La primera implementación se hizo en MATLAB 2008 las imágenes son en escala de grises con una resolución de 512 x 512, 640 x 640 y 800 x 800 la cual incluye al objeto a buscar, todo opera bajo un ambiente controlado, el objeto a buscar se encuentra dentro de una ventana de 32 x 32, 64 x 64 y 128 x 128 a continuación se muestran un ejemplo de los elementos utilizados en las pruebas:

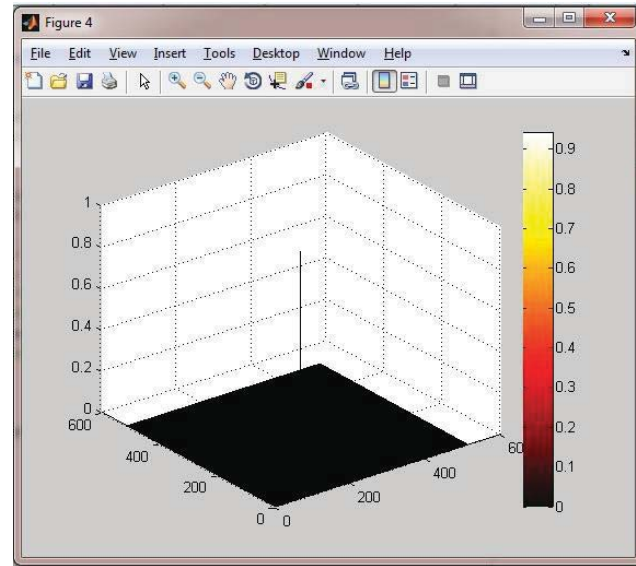


Escena de prueba



Objeto a buscar

Una vez que se le aplica el filtro obtenemos como resultado una matriz de correlación con múltiples picos debido a que todos los objetos dentro de la imagen comparten características similares, es así que se realiza un recorrido dentro de la matriz en la cual se deja únicamente el elemento con la correlación más alta obteniendo el siguiente pico:



Después de la implementación en CUDA obtuvimos los resultados mostrados en la tabla 1.

TABLA I Se muestran los tiempos promedio obtenidos en segundos, las columnas están organizadas por la resolución de la escena de prueba seguido del tamaño de filtro utilizado y las filas por la tecnología.

	Tamaño		
	512px,32px	512px,64px	512px,128px
CPU	8.4667	30.4330	117.8497
GPU	0.01131	0.03624	0.1185

	Tamaño		
	640px,32px	640px,64px	640px,128px
CPU	37.3271	140.2014	552.3482
GPU	0.01383	0.03941	0.1264

	Tamaño		
	800px,32px	800px,64px	800px,128px
CPU	58.3312	219.2611	860.7171
GPU	0.01856	0.04457	0.1292

#### IV. CONCLUSIONES Y TRABAJO FUTURO

En este artículo se presentó una implementación de una técnica para la detección de objetos utilizando una GPU con la que se pretendía agilizar el proceso, los resultados muestran un aumento considerable en cuanto a la velocidad de respuesta comparado con la implementación de la misma técnica en CPU, esto abre la posibilidad de trabajar con imágenes cada vez más grandes e inclusive realizar una implementación en tiempo real.

Actualmente se trabaja para implementar esto mismo, ahora en FPGA otra tecnología de naturaleza paralela, se pretende encontrar de entre estas 2 no solo la mejor para este problema si no también los factores que influyen para que una tenga mejor desempeño que la otra y de esta manera se puedan tomar como base para la elección de la tecnología adecuada en problemas similares.

#### REFERENCIAS

- [1] S. Martínez Díaz, 2011: "Reconocimiento robusto de imágenes: utilizando filtros de correlacion no lineal". Editorial Academia Española.
- [2] S.Martínez Díaz, 2008: Tesis doctoral "Métodos no lineales, invariantes a distorsiones geométricas e iluminación, para el reconocimiento de patrones e imágenes". CICESE.
- [3] J. Sanders, E. Kandrot, 2010: "Cuda by Example: An introduction to General-Purpose GPU programming. Addison-Wesley.
- [4] David B. Kirk, Wen-mei W. Hwu, 2012: "Massively Parallel Processors: A Hands on Approach". Morgan Kaufmann 2 edition.
- [5] MathWorks inc, 2011: "Getting Started Guide"
- [6] NVIDIA Corp.,2010 "CUDA C Programming Guide.", NVIDIA Cor., Santa Clara California. Version 3.2.
- [7] [http://ark.intel.com/es/products/75040/Intel-Core-i5-4440S-Processor-6M-Cache-up-to-3\\_30-GHz](http://ark.intel.com/es/products/75040/Intel-Core-i5-4440S-Processor-6M-Cache-up-to-3_30-GHz)

# Algoritmo paralelo sobre un sistema manycore para eliminar ruido en imágenes

Teodoro Álvarez Sánchez  
Instituto Politécnico Nacional  
CITEDI,  
Tijuana B.C, México  
talvarez@citedi.mx

Jesus G. Figueroa Nazuno  
Instituto Politécnico Nacional  
CIC,  
Ciudad México, México  
jfn@cic.mx

Jesus Álvarez Cedillo  
Instituto Politécnico Nacional  
CIDETEC,  
Ciudad México, México  
jalvarez@cidetec.mx

Roberto Charles Herrera  
Instituto Politécnico Nacional  
CITEDI,  
Tijuana B.C, México  
robcharles@citedi.mx

**Resumen**— Actualmente la computación paralela, está llamando la atención por su eficiencia de procesamiento en la ejecución de algoritmos en forma paralela de alta complejidad, dicha eficiencia de ejecución está relacionada implícitamente a la arquitectura del procesador, el orden en acceso a memoria, la organización de los procesadores (topología) en hardware y en la parte del software es importante saber la organización el sistema de cómputo, la sincronización de ejecuciones, modelos de programación, modelos de memoria, como se observara estos son los factores importantes de analizar y tomar en cuenta para obtener mayor eficiencia en el costo/rendimiento al estar trabajando con sistemas Manycore. Trabajar con un sistema manycore es importante determinar: su escalabilidad del sistema ( 1 a N cores), la capacidad de realizar ejecuciones en paralelo, concurrentes, seriales y ejecuciones en clúster de cores, además el sistema se organiza en diferentes topologías, con su respectivo repertorio de instrucciones y programas diseñados para dicha plataforma y portabilidad. Este trabajo presenta un análisis comparativo de los algoritmos ejecutados en paralelo para el filtrado y la restauración de la imagen patrón Lena con ruido gaussiano inducido, en el cual los programas integran técnicas de paralelismo para implementarse y ejecutarse en el sistema manycore Epiphany III. El experimento se realizó, para arreglos de 2x2 y 4x4 con (4, 8, 16) cores, además se establece el desempeño del sistema Epiphany en ciclos máquina, el cual se muestra los tiempos de respuesta de cada algoritmo, con bajo consumo de potencia por el número de cores utilizados para el procesamiento de la imagen.

**Palabras claves**— Ruido gaussiano; Ruido speckle; Manycore; Modelo de programación, Modelos de memoria)

## I. INTRODUCCIÓN

Hoy en día, la digitalización es profusa en diferentes áreas de aplicación comprendidas en el procesamiento de imágenes, con el objetivo de mejorar su calidad para una correcta interpretación humana o para facilitar la búsqueda de información. Hay bastantes aplicaciones que necesitan trabajar las imágenes de gran tamaño, utilizando algoritmos costosos con un alto poder de cómputo y proporcionar los resultados en un menor tiempo. Durante la formación de la imagen, adquisición, almacenamiento y/o transmisión, muchos son los factores que introducen ruido, afectando la calidad de la

imagen digital [1]. Los tipos más comunes de ruido son el ruido gaussiano, el ruido impulsivo, ruido uniforme, ruido speckle. Este artículo se verá el ruido gaussiano que aparece en el momento de la adquisición de la imagen por un sensor de una cámara, escáner u otro dispositivo en malas condiciones y tiene como consecuencia la borrosidad de todos los píxeles de la imagen, provocando un efecto de bruma (blur). Por otro lado, está el ruido impulsivo que aparece durante la transmisión de la imagen por un canal contaminado, afectando solo a ciertos píxeles de la imagen [2].

Como ya se menciona, el objetivo es el procesamiento paralelo utilizando un algoritmo paralelo para procesamiento de imágenes. En la actualidad, el uso de tecnologías para obtener un alto rendimiento computacional se ha convertido indispensable, cuando se requieren hacer intensificar el procesamiento de forma eficiente y con resultados eficaces. Varios sistemas entre ellos, clúster de computadoras y sistemas híbridos CPU-GPU ofrecen cada uno diferentes características que pueden ser aplicadas a múltiples problemas. El procesamiento de imágenes, es un área que requiere resolver aplicaciones. Por ejemplo, en las áreas de investigación científica y en los sistemas de reconocimiento biométrico. Los sistemas híbridos pretenden aprovechar la totalidad de los recursos de procesamiento computacional mediante la utilización de tarjetas gráficas, las cuales son especialmente diseñadas para soportar operaciones matriciales en paralelo, gracias al gran número de ALU's (CORES) que poseen. Esta es la razón principal que un procesamiento de imágenes sea beneficiado por el uso de este tipo de sistemas.

## II. SISTEMA MANYCORE

Recientemente, se han incorporado en el mercado de microprocesadores los llamados *sistema manycore*, que se caracterizan por aumentar aún más el poder de cálculo. Esta combinación de diferentes niveles de hardware y software, permite realizar operaciones en los *cores* de manera secuencial, concurrente o paralela en cada uno de los cores (núcleos), en estos sistemas se utiliza un esquema de un core

por hilo que tienen el acceso a la misma memoria principal. El sistema manycore es un conjunto de cores, en el que cada uno de los cores tiene un espacio de memoria compartida en el cual pueden acceder cualquier otro cores. Hay que destacar que el acceso a la memoria compartida, es un sistema homogéneo, que tiene sus limitaciones y sólo se puede acceder en forma secuencial o concurrente, sin embargo cuando se conectan los cores a través de una topología de tipo “malla”, este sistema aumenta la capacidad de acceder a la memoria de forma concurrente. Los sistemas manycore son complejos. Esto dificulta lograr un diseño óptimo que evite la sobrecarga de procesamiento y los factores que afectan sistema [12].

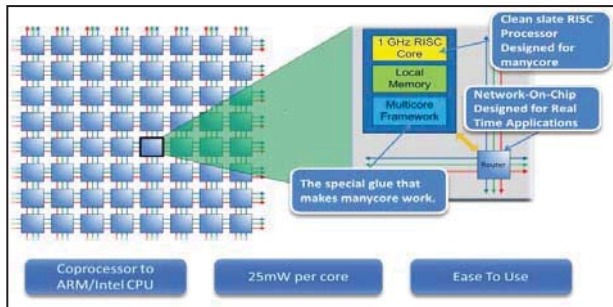


Fig. 1. Componentes de la Arquitectura Epiphany IV.

En la Fig. 1, se muestra la arquitectura Epiphany III es un reciente sistema manycore de 16 cores, entre sus virtudes, el sistema tienen un escalamiento de cores, con memoria compartida capaz de ejecutar cómputo paralelo. Este sistema ofrece un consumo de potencia de 50 Gigaflops por Watt en cálculos de simple precisión, estos cores se fabricaron en 28 nm, de tipo RISC e interconectados vía una red interna de alta velocidad de 800 MHz. Cada uno de sus cores utiliza una arquitectura similar, el cual es similar a los núcleos de los GPU o de aceleradores como Xeon Phi, obviamente sin el hardware x86 de este último [4-9]. Epiphany III utiliza una memoria plana que consta de  $2^{32}$  bytes. Las direcciones se tratan como números sin signo, desde 0 a  $(2^{32} - 1)$ . Este espacio es considerado como un conjunto de  $2^{30}$  palabras (Word) de 32 bits.

#### A. Modelo de programación paralela

En la actualidad, las computadoras modernas tienen procesadores con arquitectura manycore, las cuales integran decenas de unidades de procesamiento (core) en un sólo circuito integrado. Estos cores son capaces de ejecutar de manera paralela varias tareas con la ayuda del concepto de hilos, permitiendo con esto que las aplicaciones se ejecuten en menor tiempo. Esto sirve para explotar el paralelismo disponible a nivel de hardware y concurrente a nivel software, la aplicación debe ser descompuesta y distribuida en diferentes cores. De esta manera cada core disponible ejecuta una parte de la aplicación en paralelo junto con los otros cores. Las ejecuciones se tienen que sincronizar en las diferentes etapas de la aplicación.

El sistema Epiphany III es utilizado para el desarrollo de esta investigación, posee un modelo de programación neutral

y compatible con la mayoría de los métodos de programación paralela, incluyendo: SIMD – Single Instruction Multiple Data, SPMD – Single Program Multiple Data, programación maestro-esclavo, MIMD, Flujo de datos estático y dinámico, Array sistólico, Multihilos de memoria compartida, paso de mensajes y CSP - Procesos de Comunicación Secuencial.

#### B. Programación paralela

La multiplicación de matrices es una operación fundamental que se utiliza en muchas aplicaciones científicas. El algoritmo de la multiplicación de matrices en paralelo descrito por Sapir [17] implica la comunicación de datos entre los núcleos vecinos de acuerdo con el algoritmo de Cannon [18]. La memoria del sistema es relativamente pequeña de cada núcleo, que presenta algunos desafíos en la implementación, haciendo necesario el uso cuidadoso de espacio de memoria disponible para la comunicación entre los núcleos. A continuación se muestra cómo se puede utilizar el sistema Epiphany III para mejorar el rendimiento general de un cálculo matemático. Por simplicidad, hemos optado la multiplicación de matrices, pero los conceptos se aplican a problemas más complicados. La multiplicación de matrices se puede representar por la siguiente fórmula:

$$C_{ij} = \sum_{k=0}^{N-1} (A_{ik} B_{kj}) \quad (1)$$

Donde A y B son las matrices de entrada, C es la matriz de resultado de i y j, en cual representan la coordenada (i, j) (fila, columna), de los elementos de la matriz. El procedimiento para implementar la multiplicación de matrices en un sólo core (núcleo) se muestra a continuación:

```
for (i = 0; i < M; i++)
for (j = 0; j < M; j++)
for (k = 0; k < K; k++)
C[i][j] += A[i][k] * B[k][j]
```

Este pseudocódigo puede escribirse en lenguaje C/C++ estándar, que se compila y ejecuta en un solo núcleo, siempre que las matrices A, B y C sean colocadas en la memoria local de los núcleos. En este ejemplo de programación, no existe diferencia entre el sistema Epiphany III y alguna otra plataforma con un solo procesador ver Fig. 2. Para acelerar el cálculo matemático, es posible utilizar varios núcleos simultáneamente. Primero se distribuyen las matrices A, B, C en P tareas. Por la naturaleza estructural de una matriz, la manera más adecuada de distribuir las tareas, es por bloques pequeños.

La Fig.4, se tiene la multiplicación de matrices en cual se divide en 16 tareas, cada una de las tareas es ejecutada en cada uno de los núcleos. La transferencia de datos durante la ejecución de las tareas, entre cada nodo, se realiza mediante el mecanismo, de paso de mensajes del sistema manycore (Epiphany III) o directamente escribiendo en la memoria compartida global.

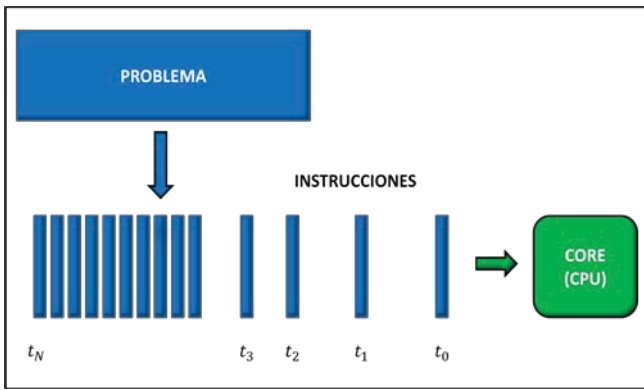


Fig. 2. Ejecución serial

Para ello se hacen programas (SPMD – Single Program Multiple Data) que se ejecutan en cada uno de los núcleos, ver Fig. 3.

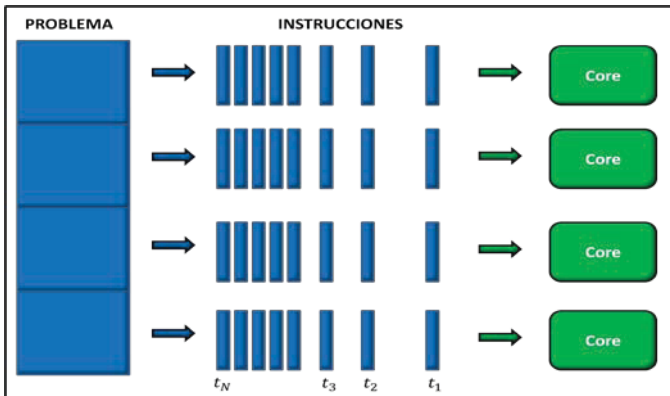


Fig. 3. Ejecución SPMD

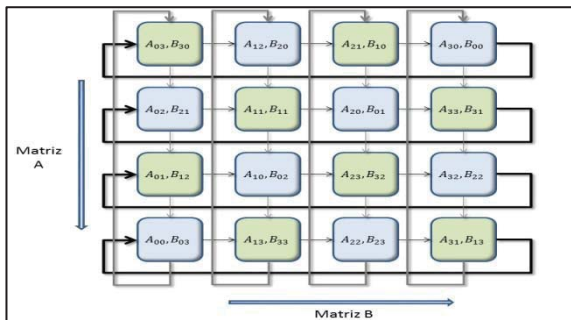


Fig. 4. Flujo de datos en la multiplicación de matrices [18]

La multiplicación paralela de matrices se realiza en cierto número de pasos determinados por la raíz cuadrática de  $P$  ( $\sqrt{P}$ ), donde  $P$  es el número de cores; con cada tarea, la multiplicación de matrices operan sobre el conjunto de datos que son del tamaño de  $\sqrt{P} \times \sqrt{P}$ . En cada paso del proceso, se hacen modificaciones a la matriz  $C$ , después la matriz  $A$  se mueve hacia abajo y la matriz  $B$  se mueve a la derecha. Este ejemplo se puede programar utilizando lenguaje de programación estándar ANSI. El sistema Epiphany III proporcionan funciones específicas para simplificar la programación manycore, pero su uso no es obligatorio, esta sistema permite a los programadores innovar en todos los

niveles. Este algoritmo, es implementado en el sistema Epiphany III con 16-cores operando a 1 GHz, resuelve la multiplicación de matrices de  $128 \times 128$  en 2 ms. Este ejemplo demuestra cómo es posible escalar el sistema para miles de cores, así como crecer linealmente el rendimiento del sistema Epiphany III, con el número de cores, cuando se utilizan modelos adecuados de programación y distribución de datos se tendrá un sistema optimizado en costo/rendimiento [10, 11,13-16,24].

### III. PROCESAMIENTO DE IMÁGENES

Una imagen digital se puede representar en dos dimensiones sobre una matriz numérica y está compuesta por un número finito de elementos, cada uno de los cuales tiene una localizado un valor, que representada por una función de intensidad de luz bidimensional  $f(x,y)$ , donde  $x$  e  $y$ , indican las coordenadas del valor de  $f$  en cualquier punto  $(x, y)$ , que es proporcional a la luminosidad (o nivel de gris) de la imagen en dicho punto. Estos elementos del arreglo digital se llaman pixeles.

Un pixel  $p$  con coordenadas  $(x, y)$  tiene cuatro vecinos de forma horizontal y vertical, cuyas coordenadas están dadas por:  $(x+1, y)$ ,  $(x-1, y)$ ,  $(x, y+1)$ ,  $(x, y-1)$ . Este conjunto de pixeles es llamado los 4 vecinos de  $p$  y se denota como  $N_4(p)$ . Los cuatro vecinos diagonales de  $p$  tienen coordenadas  $(x + 1, y + 1)$ ,  $(x + 1, y - 1)$ ,  $(x - 1, y + 1)$ ,  $(x - 1, y - 1)$  y son denotados por  $N_D(p)$ . Además se tienen otro arreglo, que se llaman los 8-vecinos de  $p$ , denotado por  $N_8(p)$ . En la Fig. 5, se muestran los pixeles vecinos del pixel central  $(x, y)$ . La adquisición de una imagen está a cargo de un dispositivo transductor, por ejemplo: sensores de una fotografía o video.

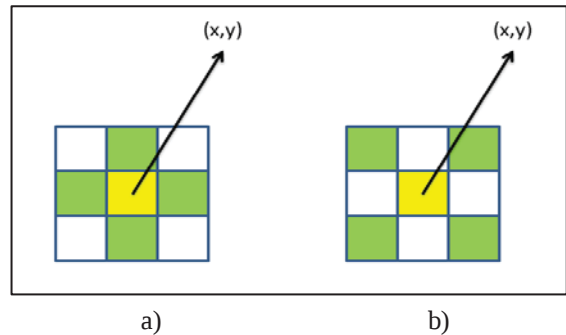


Fig. 5. a) Vecindad horizontal y vertical del pixel  $(x, y)$ , b) Vecindad diagonal del pixel  $(x, y)$ .

Es importante señalar que durante esta etapa, estos transductores de imágenes agregan ruido a la imagen degradando su calidad. Una imagen está contaminada cuando algunos de sus pixeles han sido alterados o transformados, perdiendo la información original de  $B/N$  o color. Esto puede deberse a varios factores, entre los que se encuentran las condiciones ambientales durante la adquisición, o por la calidad propios de los dispositivos transductores, otro ejemplo, en una cámara CCD los niveles de luz y temperatura de los sensores, además del movimiento de la cámara o



desenfoque son los principales factores que afectan la cantidad de ruido en la imagen resultante. Entre los ruidos más comunes de funciones de densidad probabilística [20] están los siguientes: ruido gaussiano: ruido uniforme, ruido speckle, ruido Impulsivo.

Una imagen que contiene ruido se elimina mediante la aplicación de filtros para eliminación de ruidos. Si inicialmente tenemos una imagen con ruido ( $I_R$ ), la cual es restaurada para eliminar el ruido, y finalmente obtener una imagen filtrada ( $I_F$ ). El filtrado de ruido, es un proceso para discriminar entre la información original de la imagen y los datos ruidosos, que impiden la interpretación correcta de la información. La etapa de filtrado de la imagen es probablemente la tarea más común en el procesamiento digital de imágenes, detección de bordes, reconocimiento de patrones, etc, es un paso fundamental en los sistemas de visión por computadora en el que las imágenes se utilizan para análisis automático o inspección humana [21].

Existen varios filtros y métodos de detección para reducir o eliminar el ruido que presente una imagen digital. Los filtros pueden ser basados en el dominio espacial o en el dominio de la frecuencia. Los filtros basados en el dominio de la frecuencia realizan operaciones sobre la transformada de *Fourier* de una imagen. Por otro lado, los métodos basados en el dominio espacial, se realizan sobre los píxeles de la imagen.

Los filtros en el dominio de la frecuencia consisten en modificar la transformada de *Fourier* para lograr un objetivo específico y después calcular la transformada inversa para obtener el resultado procesado. La transformada de *Fourier* se ve como una transformación del dominio espacial al dominio de la frecuencia sin perder información de la imagen. Se pueden mencionar los filtros de paso bajo y de paso alto que se usan para atenuar las bajas y altas frecuencias respectivamente. Los filtros en el dominio espacial consisten en una operación sobre los píxeles de la vecindad. El filtrado crea un nuevo pixel con coordenadas iguales a las coordenadas del centro de la vecindad y cuyo valor es el resultado de la operación de filtrado. Si la operación se realiza sobre los píxeles de la imagen, el filtrado es espacial lineal, en caso contrario es no lineal.

En este trabajo de investigación se ha utilizado la transformada rápida de *Fourier*, basado en el dominio de la frecuencia, para mejorar la calidad de la imagen y eliminar el ruido gaussiano en la imagen de prueba.

#### A. Dominio frecuencial

Una función senoidal de dos dimensiones, se caracteriza por su fase, su frecuencia de oscilación y su dirección de oscilación. Una función senoidal en un espacio de dos dimensiones, que se expresa a continuación:

$$f(m, n) = \sin\{2\pi(Um + Vn)\}. \quad (2)$$

Donde  $m, n$  son las coordenadas espaciales (en píxeles),  $U$  y  $V$  son las dos frecuencias (ciclos/pixel). El argumento del seno son los radianes. La frecuencia y la dirección de máxima oscilación son:

$$\Omega = \sqrt{U^2 + V^2}, \quad \theta = \arctan\left\{\frac{V}{U}\right\}. \quad (3)$$

Se suele trabajar con las frecuencias normalizadas al tamaño de la imagen  $M \times N$ .

$$(u, v) = (MU, NV). \quad (4)$$

Donde ahora las dimensiones son ciclo/imagen. La expresión de la sinoidal es ahora:

$$f(m, n) = \sin\left\{2\pi\left(\frac{u}{M}m + \frac{v}{N}n\right)\right\}. \quad (5)$$

Las frecuencias  $U, V$  tienen un periodo 1. Esto quiere decir que se obtiene la misma imagen para  $U = 0.2$  que para  $U = 1.2$ . Aplicando que:

$$\sin(\alpha + \beta) = \sin(\alpha) \cos(\beta) + \cos(\alpha) \sin(\beta). \quad (6)$$

Una señal se puede descomponer en una sumatoria de funciones de seno (armónicos) con diferentes frecuencias, la transformada de *Fourier* nos da la proporción en que contribuye cada una de ellas. Por ejemplo, en la Fig. 6, se tiene una señal senoidal  $y(x)$  y la magnitud de la transformada de *Fourier* de dicha señal.

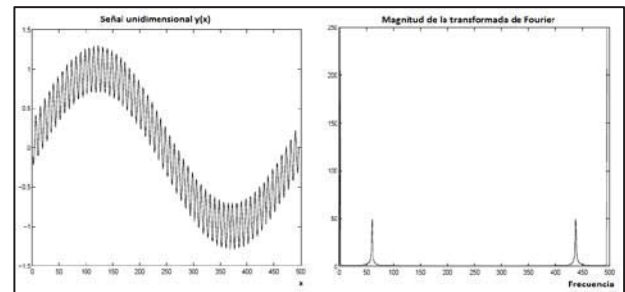


Fig. 6. a) Señal  $y(x)$ ; b) Magnitud de la transformada de *Fourier*.

#### B. Transformada de *fourier* en dos dimensiones

La transformada de *Fourier* para una imagen de tamaño finito  $M \times N$  vendrá dada por:

$$F(U, V) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) e^{-j2\pi(Um+Vn)}. \quad (7)$$

La transformada de *Fourier*  $F(U, V)$  es una función continua que se tienen que discretizar. Su transformada inversa:

$$f(m, n) = \int_0^1 \int_0^1 F(U, V) e^{j2\pi(Um+Vn)} dU dV. \quad (8)$$

Haciendo un cambio de variable:  $U = \frac{u}{M}$ ,  $V = \frac{v}{N}$ .

$$F(u, v) = \frac{1}{MN} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) e^{-j2\pi(\frac{u}{M}m + \frac{v}{N}n)}. \quad (9)$$

La Transformada de Fourier inversa en forma discreta sería:

$$f(m, n) = \frac{1}{MN} \sum_{u=0}^{M-1} \sum_{v=0}^{N-1} F(u, v) e^{j2\pi(\frac{u}{M}m + \frac{v}{N}n)}. \quad (10)$$

Donde ahora los índices van desde  $0 < u < M-1$ ,  $0 < v < N-1$ .

$$F(u, v) = |F(U, V)|. \quad (11)$$

Se sabe que  $U$  puede tener estos valores de  $0 < U < 1$ , entonces también los valores  $M$ , es decir que  $u = UM$ , que se obtendrá el valor de  $u = 1, \dots, M$ .

Por lo tanto, se tiene que la Transformada de Fourier se puede hacer una transformación de un dominio (en el espacio) a otro dominio (en frecuencia) sin perder información de la imagen. El representar la información de la imagen en el dominio de la frecuencia el cual tiene ventajas, a la hora de aplicar algunos algoritmos que de determinar ciertas propiedades de la imagen. Cada armónico va a recoger un nivel de detalle, de cómo es la variación de los niveles de gris especialmente.

$$F(U, V) = R(U, V) + jI(U, V). \quad (12)$$

Donde,

$$R(U, V) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \cos 2\pi(Um + Vn). \quad (13)$$

$$I = (U, V) = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f(m, n) \sin 2\pi(Um + Vn). \quad (14)$$

Los algoritmos se aplican sobre el espectro de magnitud,

$$|F(U, V)| = \sqrt{R^2(U, V) + I^2(U, V)}. \quad (15)$$

La fase se calcula con la siguiente expresión:

$$\text{ang}(F(U, V)) = \arctan \frac{I(U, V)}{R(U, V)}. \quad (16)$$

Normalmente no se usa la información de la fase en el análisis de imagen, porque suele tener varios aleatorios, aunque su información es crucial para reconstruir la imagen original. Como la transformada de Fourier es una función periódica, un periodo de  $M \times N$  es suficiente para representarla. Se recuerda que  $M$  y  $N$  representan el tamaño de la imagen.

Si construimos una imagen cuyo nivel de gris sea proporcional a la magnitud de la transformada de Fourier, podemos representar gráficamente a la transformada de Fourier. Las características que podemos obtener de una imagen representada por su espectro de frecuencia se agrupan por zonas; a continuación se describen estas características:

- Las zonas homogéneas en la imagen darán lugar a que la energía del espectro esté concentrada en las bajas frecuencias.
- Zonas con muchos bordes, transiciones frecuentes de los niveles de gris dará lugar a un espectro con componentes de alta frecuencia.
- Si en la imagen existe alguna regularidad, es decir, patrones que se repiten, dará lugar a picos de igual intensidad separados una distancia regular.

A continuación se tiene la expresión que mejorara la visualización de la imagen  $\log_2(1 - |F(u, v)|)$ . En la Fig. 7, se muestra las intensidades de tono.

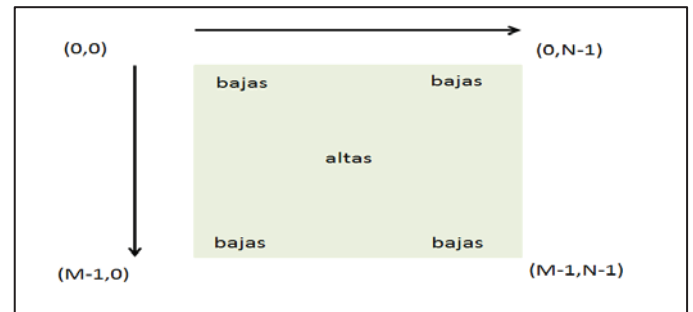


Fig. 7. Visualización de la transformada discreta de Fourier

#### IV. IMPLEMENTACIÓN DEL ALGORITMO PARALELO EN EPIPHANY

Para explotar el paralelismo disponible en alguna plataforma en hardware, se sigue que la aplicación debe ser descompuesta en diferentes porciones. De esta manera cada una de los Core podría ejecutar una parte del programa en paralelo con otros Cores. La descomposición de la aplicación debe ser seguida por una sincronización de las diferentes partes para asegurar la consistencia de los resultados.

##### A. Descomposición de tareas y datos

Dependiendo de la naturaleza del problema, podemos elegir entre dos enfoques de las descomposiciones: la descomposición de datos o la descomposición de tareas.

**Descomposición de datos:** En el enfoque de la descomposición de datos, deben ser procesados, y divididos en partes. Las mismas instrucciones, se pueden ejecutar en paralelo en diferentes núcleos. Este enfoque es adecuado cuando las partes de datos pueden ser procesados en forma independiente. Con  $n$  cores es posible resolver el problema  $n$  veces más rápido que haciendo uso de un solo núcleo. Salvo por el mínimo retraso que supone el reparto inicial de la carga de trabajo, y la recolección final de los datos, el cual se obtiene una aceleración lineal con el número de procesadores.

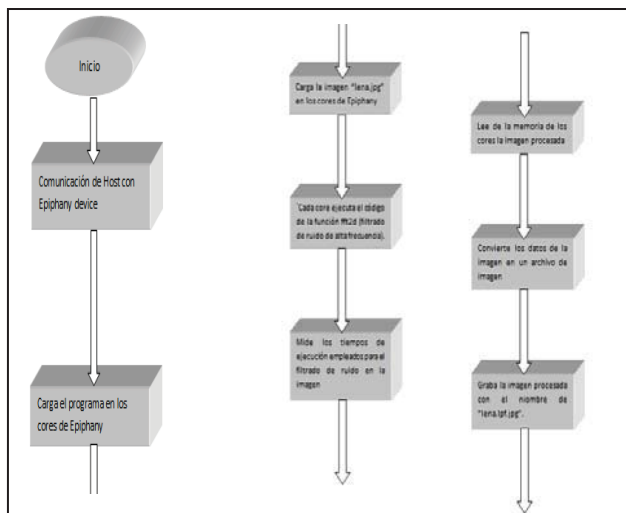
**Descomposición de tareas:** En esta descomposición de tareas, todo el problema se debe dividir en tareas. Una tarea es una secuencia de programa que se puede ejecutar de forma independiente, en paralelo con otras tareas. Este enfoque es beneficioso cuando las tareas mantienen altos niveles de independencia.

Los algoritmos de cada tarea son diferentes, este modelo se denomina descomposición funcional y se puede aprovechar las características particulares de cada tipo de tarea para ejecutarlas en el núcleo que sea más conveniente para tal efecto.

V. ALGORITMO PARA EJECUTARSE EN PARALELO

En este programa se demuestra el procedimiento de la ejecución para el procesamiento de imágenes. La imagen contiene ruido adicionado y puesto en el archivo "lena.jpg". El programa carga esta imagen de 128 x128 pixeles y la filtra mediante un filtro de alta frecuencia, generando un archivo de imagen "lena.lpf.jpg". El programa utiliza una biblioteca, *DevIL library* para leer y cargar la imagen desde donde se encuentra almacenada. Después de leer el archivo y comprobar que tiene el tamaño correcto, los pixeles de la imagen son cargados en los cores del sistema Epiphany III, donde son procesados mediante la aplicación del filtro de alta frecuencia. El diagrama de bloques de la Fig. 8, muestra los pasos del algoritmo aplicado a la imagen.

EL programa termina creando una imagen sin ruido, llamada "lena.lpf.jpg" de 128x128 pixeles. La plataforma manycore de Epiphany III se adapta fácilmente a la implementación del algoritmo como una aplicación SIMD, donde todos los cores ejecutan el mismo código sobre datos únicos (porción de la imagen) [22,23].



a)

```

Loading original image from file ".../lena.jpg".
Width: 128 Height: 128 Depth: 1 Bpp: 24
Writing image to Epiphany
.....
GO!
Done!
Finished calculation in 593200 cycles (0.989 msec @ 600 MHz)
FFT2D - 284435 cycles (0.474 msec)
FFT Setup - 189 cycles (0.000 msec)
BITREV - 8996 cycles (0.015 msec)
FFT1D - 127230 cycles (0.212 msec x2)
Corner Turn - 4491 cycles (0.007 msec)
LPF - 22225 cycles (0.037 msec)
Reading processed image back to host
.....
Saving processed image to file ".../lena.lpf.jpg".
    
```

b)

Fig. 8. a) Diagrama de flujo, b) tiempo procesamiento.

En esta aplicación todos los cores se sincronizan para operar sobre la memoria compartida, permitiendo un tránsito fácil de datos entre cores. El sistema tiene 16 cores y cada core tiene 32KB de memoria local. El tamaño de la imagen ha sido limitado a 16 KB (128 x128 pixeles), justo el espacio de memoria.

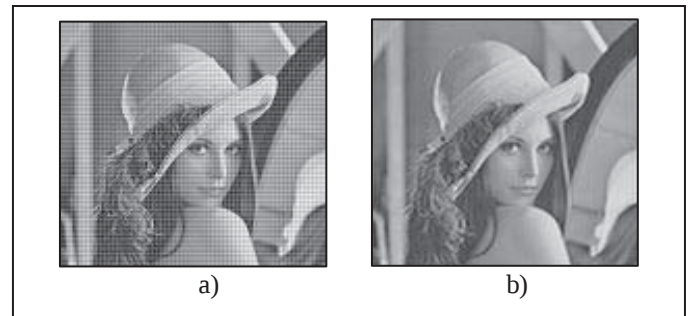


Fig. 9. Imagen "lena" a) imagen con ruido b) imagen filtrada con fft\_2d utilizando epiphany.

VI. RESULTADOS

El algoritmo paralelo se ejecutó correctamente utilizando arreglos de 2x2 y 4x4 en los cores. La comparación basada en los tiempos de ejecución del algoritmo, se muestra en la Fig. 10.

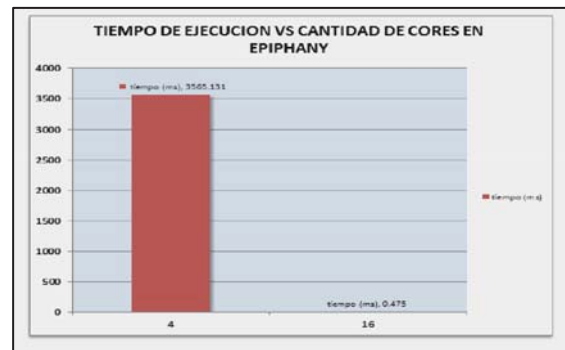


Fig. 10. Desempeño de Epiphany III con arreglos de 4 y 16 cores.

La frecuencia de reloj fue utilizada de 600 MHz para cada Core; la cantidad de ciclos de ejecución para cada una de los algoritmos se muestra en la Tabla 1.

TABLA 1.

CANTIDAD DE CORES	2x2 CORES 4	4x4 CORES 16
FFT2D	2'139,078,656	284,435
FFT setup	1'935,642,304	189
BITREV	4'293,918,720	8996
FFT1D	2'360,365,376	127,230
Corner turn	4'294,959,104	4465
LPF	24,576	22,225

Nota: Todos los tiempos son expresados en ciclos de reloj (600 MHz)

Para el procesamiento de imágenes se realizó a una frecuencia de 600 MHz. Las pruebas realizadas sobre el mismo algoritmo para un arreglo de cores de 2x2 y 4x4, se muestran una imagen del sistema Epihaphny III en la Fig. 11.



Fig. 11. Sistema Epiphany III con 16 cores

Aunque el tiempo de ejecución del algoritmo queda delimitado por el tamaño de la imagen, resulta interesante el análisis comparativo del tiempo entre ambos arreglos de cores, por ello podemos resaltar las siguientes características importantes de este trabajo:

#### CONCLUSIONES

- ✓ El algoritmo de la transformada rápida de Fourier (FFT) se aplica a una imagen contaminada con ruido gaussiano. Este algoritmo se prueba sobre arreglos de cores de 2, 4, 16 ejecutados paralelamente.
- ✓ Al aplicarse sobre un core (ejecución secuencial), resulta irrelevante la comparación del tiempo de respuesta de la ejecución.
- ✓ Se reduce el tiempo de ejecución considerablemente al aplicarlo sobre 2, 4, 16 cores.
- ✓ De esta manera se aprovecha la arquitectura manycore para ejecutar gran demanda de datos para el cálculo matemático requerido en el procesamiento de imágenes.
- ✓ Aunque el tiempo de ejecución varía de acuerdo al tamaño de la imagen, éste queda limitado por la capacidad de memoria interna del sistema, es importante mencionar que el tiempo de ejecución se incrementará sobre imágenes de mayor tamaño.

#### Agradecimientos

El resultado de este trabajo es parte de los apoyos recibidos de la Secretaria de Investigacion y Posgrado del IPN para el proyecto SIP-20160693

#### REFERENCIAS

- [1] González & Woods, "Digital Image Processing", 3ra. ed.
- [2] Camarena, "Progress in Pattern Recognition, Image Analysis, Computer Vision", 2009.
- [3] Grama A, Gupta A, Karypis G, Kumar V, "An Introduction to Parallel Computing Design and Analysis of Algorithms", 2nd ed., Pearson Addison Wesley, 2003.
- [4] Burger T, "Intel Multi-Core Processors: Quick Reference Guide", (2014, Marzo 16) [En línea]. Disponible en: [http://cachewww.intel.com/cd/00/00/23/19/231912\\_231912.pdf](http://cachewww.intel.com/cd/00/00/23/19/231912_231912.pdf).
- [5] D. Wentzlaff, P. Griffin, H. Hoffmann, L. Bao, B. Edwards, C. Ramey, M. Mattina, C. C. Miao, J. F. B. III, and A. Agarwal, "On-chip interconnection architecture of the tile processor", IEEE Micro, vol. 27, pp. 15-31, 2007.
- [6] Kalray, "Kalray MPPA 256", User's Guide, (2014, Febrero 6) [En línea]. Disponible en: <http://www.kalray.eu/products/mppa-manycore/mppa-256/>.
- [7] Adapteva, "The Parallella Board", User's Guide, (2014, Febrero 6), " [En línea]. Disponible en: <http://www.adapteva.com/parallella-board/>.
- [8] Intel, "Intel Xeon Phi Coprocessor", User's Guide, (2014, Febrero 6) [En línea]. Disponible en: <http://www.ssl.intel.com/content/dam/www/public/us/en/documents/datasheets/xeonphi-coprocessor-datasheet.pdf>.
- [9] Intel, "The single-chip-cloud computer", (2014, Febrero 6) User's Guide, [En línea]. Disponible en: <http://www.intel.com/content/dam/www/public/us/en/documents/technologybriefs/intel-labs-single-chip-cloud-rticle.pdf>
- [10] L. Benini and G. De Micheli, "Networks on chips: a new soc paradigm", Computer, vol. 35, pp. 70-78, January 2002.
- [11] Isaac D Scherson and Abdou S. Youssef, "Interconnection Networks for High Performance Parallel Computers", IEEE Computer Society Press, 1994
- [12] Teodoro Álvarez Sánchez & Miguel Ángel Miranda Romeo, "Factors influencing manycore processor".
- [13] Wittwer Tobias, "An Introduction to Parallel Programming", VSSD, Leeghwaterstraat 42, 2628 CA Delft, The Netherlands, (First edition 2006).
- [14] J. Dongarra et al, "Sourcebook of Parallel Computing", Morgan Kaufmann Publishers, 2003.
- [15] K. Hwang, Z. Xu, "Scalable Parallel Computing", McGraw Hill, 1998.
- [16] Yaniv Sapir, "Scalable Parallel Multiplication of Big Matrices", (2014, Febrero 6) [En línea]. Disponible en: <http://www.adapteva.com/whitepapers/scalable-parallel-multiplication-of-big-matrices/>.
- [17] L. E. Cannon, "A cellular computer to implement the Kalman filter algorithm", DTIC Document, Tech. Rep., 1969.
- [18] Adapteva, User's Guide, "Epiphany Technical Reference Documents", (2014, Febrero 6) [En línea]. Disponible en: <http://www.adapteva.com/all-documents>.
- [19] González, R. y R. Woods, "Digital Image Processing", Person Education. 2008.
- [20] Camarena, J. G. y col, "Fast Detection and Removal of Impulsive Noise using Peer Group and Fuzzy Metrics", 2008.
- [21] Plataniotis, K.N. y A.N. Venetsanopoulos, "Color Image Processing and Applications", 2000.
- [22] A. Vajda, "Programming Many-Core Chips", Springer, 2011.
- [23] J. Díaz, C. Muñoz-Caro, and A. Nino, "A survey of parallel programming models and tools in the multi and many-core era", Parallel and Distributed Systems, IEEE Transactions on, 23(8):1369–1386, Aug. 2012.
- [24] Hugué, Dr. Michelle, User's Guide, (2014, Febrero 6) [En línea]. Disponible en: <http://cmisc411.com/distributed-memory-architecture>.

# Un algoritmo memético paralelo para TSP

Joel Artemio Morales Viscaya  
 División de Estudios de Posgrado e Investigación  
 Instituto Tecnológico de La Paz  
 La Paz, B.C.S., México  
 Email: iscviscaya@gmail.com

Marco Antonio Castro Liera  
 División de Estudios de Posgrado e Investigación  
 Instituto Tecnológico de La Paz  
 La Paz, B.C.S., México  
 Email: mcastro@itlp.edu.mx

**Resumen**—Se presenta un nuevo algoritmo memético aplicado al problema del agente viajero, con una estrategia de búsqueda local que mezcla características de 3-OPT y Lin-Kernighan. El algoritmo se implementó en un cluster de computadoras y se incluyen resultados estadísticos de su desempeño al aplicarlo en diferentes instancias obtenidas de la TSPLIB de entre 100 y 300 ciudades. Además se demuestra su rendimiento superior al de otras heurísticas como ACO, AG, 2-OPT y 3-OPT.

## I. INTRODUCCIÓN

El problema del agente viajero simétrico (TSP por sus siglas en inglés, Traveling Salesman Problem) es un problema clásico y ampliamente estudiado de optimización combinatoria que consiste en encontrar, dado un conjunto de ciudades, la ruta más corta que permita a un agente viajero visitar todas ellas y regresar a la ciudad en la que inició el recorrido, en donde la distancia entre las ciudades debe ser simétrica; es decir, la distancia entre la ciudad A y la ciudad B debe ser la misma que entre B y A para cualquier par de ciudades en el conjunto.

TSP pertenece a la clase de problemas de optimización NP-duro (es un problema duro dentro de los no polinomiales) de acuerdo a [1] y es importante, pues constituye un componente central de un número considerable de problemas del mundo real. Sus aplicaciones incluyen pero no están restringidas a la logística, la fabricación de circuitos integrados y tableros de circuitos perforados, medición por rayos x, data clustering, mira de telescopios, análisis de secuencias genéticas, guiado de láser para cristal, conexión de antenas, construcción, programación de tareas, la atención a llamadas de emergencia, servicio postal, entre otras.

TSP es posiblemente el problema más famoso y extensamente estudiado en el campo de la optimización combinatoria [2]. En años recientes, muchas heurísticas o meta-heurísticas han sido aplicadas para resolver problemas NP-duros (ya que por definición no existen algoritmos que los resuelvan en tiempos razonables), como por ejemplo el recocido simulado [3], la Búsqueda Tabú (TS) [4], Algoritmos Genéticos[5], Redes Neuronales (RN) [6], Optimización por Colonia de Hormigas (ACO) [7], Optimización por enjambre de partículas (PSO) [8], entre otras.

Los algoritmos existentes para resolver TSP se pueden dividir en dos clases: algoritmos exactos y algoritmos de aproximación. Los algoritmos exactos pueden garantizar que convergen a la solución óptima en una cantidad acotada de pasos. De momento, el algoritmo exacto más eficiente

está basado en el corte de planos y técnicas de programación lineal [9] y ha sido aplicado a instancias grandes de TSP. Sin embargo, estos algoritmos exactos tienen una complejidad computacional muy elevada [10], además de ser muy difíciles de programar y requerir grandes cantidades de almacenamiento en memoria. En contraste, los algoritmos de aproximación obtienen buenas soluciones en tiempos razonables, pero no pueden garantizar que van a converger a soluciones óptimas. Los algoritmos de aproximación para TSP se pueden dividir a su vez en dos categorías: algoritmos de construcción de rutas [11] y algoritmos de mejoramiento de rutas [12].

Los algoritmos de construcción de rutas van construyendo la ruta (o tour) de manera gradual, añadiendo una nueva ciudad en cada paso, mientras que los algoritmos de mejoramiento de rutas empiezan con una solución inicial y tratan de mejorarla en cada paso, realizando intercambios.

## II. ALGORITMOS MEMÉTICOS

Los algoritmos evolutivos (EA) pertenecen a una rama de la inteligencia artificial que involucra problemas de optimización y se inspiran en los mecanismos de la evolución biológica. Son un tipo de estrategia de búsqueda global [13] que se ha utilizado con éxito para resolver múltiples problemas [14]. Sin embargo, en particular para problemas combinatorios complejos se considera que las estrategias puras de este tipo, como son los algoritmos genéticos, no son suficientemente buenas [15].

---

### Algoritmo 1 Estructura general de un algoritmo memético

---

```

1: t=0
2: P(t)=InicializarLaPoblación
3: P(t)=BúsquedaLocal(P(t))
4: P(t)=Selección(P(t))
5: mientras criterio de parada no cumplido haz
6:   P(t)=Cruza(P(t))
7:   P(t)=Mutación(P(t))
8:   P(t)=BusquedaLocal(P(t))
9:   P(t+1)=Selección(P(t))
10:  t=t+1
11: fin mientras

```

---

La combinación de algoritmos evolutivos poblacionales con heurísticas de búsqueda local forman lo que se conoce como un algoritmo memético (MA). Debido a la mezcla de las

propiedades incluyentes de los EA y la exploración a detalle de zonas prometedoras de la búsqueda local, los MA han mostrado ser más eficientes y eficaces en la resolución de problemas combinatorios [16], [17]

El pseudocódigo general de un MA basado en un AG se puede observar en el algoritmo 1.

### III. ALGORITMOS DE BÚSQUEDA LOCAL

Un enfoque heurístico para los problemas de optimización combinatoria es la mejora iterativa de las soluciones, mediante una transformación que tome soluciones del problema y las lleve a nuevas soluciones con mejor aptitud, hasta que ya no sea posible encontrar dichas transformaciones.

Para TSP se han estudiado algoritmos de búsqueda local conocidos como  $k$ -opt, cuya idea central es tomar un recorrido hamiltoniano (un tour que visita todas las ciudades y regresa a la inicial) y llevar a cabo el reemplazo de  $k$  aristas o caminos en el recorrido por  $k$  aristas que no se encuentren en el mismo con la condición de seguir formando un recorrido hamiltoniano, de acuerdo a [13] se considera que para  $k > 4$  son demasiado costosas.

### IV. CLÚSTER DE COMPUTADORAS

Una alternativa para disminuir el tiempo de ejecución en la resolución de problemas complejos de ingeniería es el cómputo paralelo, que consiste en la ejecución de múltiples instrucciones de un programa simultáneamente en varias unidades de procesamiento [18], [19], una forma de implementar el cómputo paralelo son los clústers de computadoras, que consisten en un conjunto de equipos de cómputo interconectados que colaborativamente intercambian y procesan información [20]. Sólo los procesadores locales tienen acceso directamente a la memoria local en cada nodo, por tanto, si un nodo requiere acceder a la memoria de otro nodo, se hace con un modelo de comunicación de paso de mensajes. Dentro del conjunto de nodos a uno de ellos se le asigna la tarea de iniciar, distribuir tareas y recopilar resultados de ejecución de la aplicación paralela, este nodo se denomina maestro [20], como se muestra en la Fig. 1

MPI es una estandarización de las librerías de paso de mensajes que fue creada por el MPI-FORUM, es portable y tiene librerías para ser implementado en lenguaje C, C++, Fortran, Java, Python y Perl, entre otros.

Existen varias implementaciones del estándar, algunas de ellas de software libre como OpenMPI.

### V. ESTRATEGIA PROPUESTA

#### V-A. Estructura

Se propone un algoritmo que se puede dividir en tres fases, la inicialización de los procesos, uno maestro y los restantes esclavos. El núcleo o la optimización, que es un algoritmo memético basado en un algoritmo genético de subpoblaciones estáticas con migración y una última fase en que cada proceso esclavo envía su mejor resultado al proceso maestro, para que encuentre el mejor de todos.

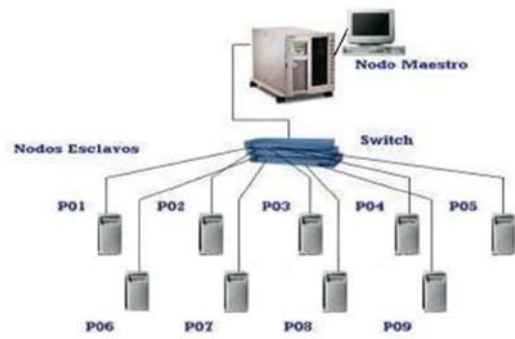


Figura 1. Esquema de un clúster de computadoras

Los procesos esclavos actúan sobre cada una de las subpoblaciones. Una vez que se cumple un cierto número de generaciones en el algoritmo memético, cada nodo esclavo comparte información con los demás en un símil a la migración. Esta estrategia es lo que se conoce en AG como modelo de islas y ha demostrado producir mejores soluciones que poblaciones aisladas o una sola población masiva [21].

Una vez que se cumple la condición de parada (el número de generaciones o vueltas del algoritmo) cada nodo esclavo envía sus resultados hacia el nodo maestro.

#### V-B. Algoritmo memético

Para garantizar una alta diversidad en las soluciones iniciales de nuestro algoritmo memético, empezamos generando una población inicial de soluciones de forma aleatoria. Sobre dichas soluciones se aplica el algoritmo de búsqueda local.

Un AG consta de tres operaciones básicas que se conocen como operadores genéticos: selección, cruce y mutación. Selección es el proceso por el cual se elige que individuos o soluciones permanecieran en la generación siguiente. Decidimos utilizar elitismo, que consiste en simplemente conservar a los mejores individuos de la población.

Cruza es como se le conoce al proceso que genera nuevas soluciones utilizando los individuos que superaron la etapa de selección, se decidió utilizar un algoritmo conocido como Order Crossover (OX), propuesto por Davis [22] para permutaciones.

La mutación se considera como un operador secundario en los algoritmos genéticos. Se aplica solamente a una cantidad reducida de los individuos y produce cambios aleatorios grandes en sus elementos.

Además, cada cierto número de generaciones se envía una cantidad prefijada de individuos (los mejores) hacia otra subpoblación. Esta operación se conoce como migración.

Se modificó el ciclo principal de un AG, agregando después de los operadores genéticos de cruce y mutación el algoritmo de búsqueda local, como se puede observar en la Fig. 2

#### V-C. Búsqueda local

Debido a que las estrategias puras 2-opt y 3-opt mejoran las soluciones de manera muy lenta y que la más utilizada Lin-Kernighan (LK), basada en determinar cuántas aristas

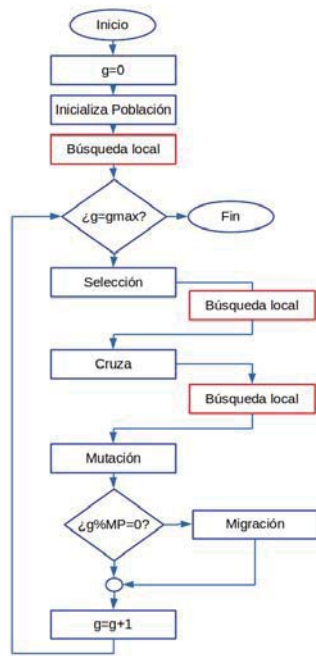


Figura 2. Diagrama del algoritmo memético propuesto

cambiar en cada iteración es computacionalmente demasiado costosa y difícil de programar, además de requerir mucho almacenamiento en memoria, se propone un algoritmo de mejoramiento de rutas nuevo (búsqueda local) basado en 3-opt y en la estrategia LK.

Ya que se deben visitar todas las ciudades, TSP se puede reducir a encontrar el orden en que deben visitarse las mismas, lo cuál hace que la manera más adecuada de representar soluciones sea como un vector de  $n$  elementos, con  $n$  igual al número de ciudades de la instancia.

Sin embargo, para aplicar los algoritmos de búsqueda local es más adecuado representar el problema en términos de grafos, vértices y aristas, por lo que es necesario primeramente traducir estas soluciones a un conjunto de aristas.

El algoritmo que se desarrolló se basa en hacer reemplazos de tres aristas, pero eligiendo qué aristas se van a reemplazar y, por cuáles aristas van a ser reemplazadas de manera “inteligente”, de manera similar a cómo ocurre en LK. A grandes rasgos, el algoritmo consta de los siguientes pasos:

1. Se elige al azar un nodo  $n_1$  contenido en la ruta original.
2. Se selecciona cualquier arista  $x_1$  adyacente a  $n_1$ .
3. Se elige una arista  $y_1$  que conecte el otro extremo de  $n_1$  tal que  $y_1$  tenga un costo menor a  $x_1$ . Si no es posible encontrar  $y_1$ , se regresa a 2 y se selecciona la otra arista adyacente a  $n_1$ , si tampoco se puede encontrar  $y_1$  se selecciona otro nodo  $n_1$ .
4. Se elige una arista  $x_2$  que contiene al extremo de  $y_1$  que no está en  $x_1$ , dando prioridad a la de costo mayor.
5. Se elige de manera aleatoria  $x_3$ , cómo cualquier arista contenida en la ruta original distinta de  $x_1$  y  $x_2$ .

6. Se eligen  $y_2, y_3$  tales que al reemplazar con  $x_1, x_2, x_3$  por  $y_1, y_2, y_3$  el recorrido resultante sea hamiltoniano (pueden no ser únicas).
7. Si algún intercambio encontrado en 6 produce un costo menor al inicial, se efectúa el reemplazo y se regresa al paso 1.
8. En caso contrario, si ningún par  $y_2, y_3$  produce un recorrido mejor, regresamos a 5 y seleccionamos otra  $x_3$ .
9. En caso de haber probado todos los nodos  $x_3$  posibles, se regresa al paso 4 y se elige una nueva  $x_2$ .
10. En caso de haber probado todos los  $x_2$  posibles, se regresa al paso 3 seleccionando una arista  $y_1$  distinta. Si ya se probó con todas las aristas  $y_1$ , entonces se ha encontrado un óptimo local y el algoritmo termina.

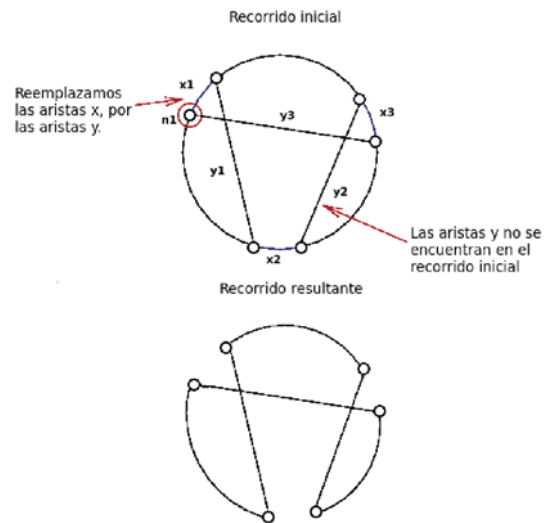


Figura 3. Ejemplo de iteración posible con reemplazo de tres aristas

Una posible iteración se puede observar en la Fig. 3

## VI. IMPLEMENTACIÓN Y PRUEBAS

Para el desarrollo del algoritmo memético se utilizó lenguaje C, compilado con GCC 4.8.4, un conjunto libre de compiladores distribuidos bajo la licencia GPL, se utilizó la librería OpenMPI en su versión 1.6.4 para el paso de mensajes entre los procesos, sobre Linux Mint 17.3 de 64 bits.

Se eligieron 20 instancias de tamaño mediano (entre 100 y 300 ciudades) de TSP para llevar a cabo las pruebas, dichas instancias se obtuvieron de la librería TSPLIB creada por el alemán G. Reinelt [23].

Las pruebas fueron llevadas a cabo en un clúster de 4 computadoras con procesadores i7-3770 con 8 GB de memoria RAM.

Se decidió hacer la comparación contra las estrategias heurística más utilizadas para resolver este problema, la optimización por colonia de hormigas, concebida originalmente para resolver TSP y un AG puro, el más utilizado de todos los algoritmos heurísticos de optimización. Debido a la naturaleza estocástica de los algoritmos heurísticos, se decidió utilizar una

muestra de 100 ejecuciones con cada una de las estrategias heurísticas.

De igual forma se comparó el desempeño de la estrategia memética con el algoritmo de búsqueda local desarrollado, contra el desempeño de la estrategia utilizando 2-opt y 3-opt.

## VII. RESULTADOS EXPERIMENTALES

En la tabla I se presenta una comparación entre el algoritmo de búsqueda local propuesta, contra una versión acelerada de 2-opt y contra 3-opt al utilizarlos para aproximar soluciones de la instancia de 225 ciudades propuestas por Reinelt de la TSPLIB, cuyo tour óptimo tiene una distancia de 3919.

Tabla I  
COMPARACIÓN DE ALGORITMOS DE BÚSQUEDA LOCAL

Algoritmo	Distancia	Varianza	Tiempo	ERP
2-opt	4319.47	75.79	1.16s	10.21 %
3-opt	4053.39	33.57	8.27s	3.42 %
Propuesta	4044.68	45.64	0.60s	3.20 %

En la tabla II se muestra una comparación entre sistema de hormigas (AS), un algoritmo genético (AG) y el algoritmo memético propuesto (AM) al resolver la instancia de 100 ciudades propuestas por Krolak/Felts/Nelson de la TSPLIB. Dicha instancia tiene un tour óptimo con un costo de 21282. Las tablas III, IV y V muestran los parámetros para cada heurística.

Tabla II  
COMPARACIÓN DE HEURÍSTICAS

Algoritmo	Distancia	Tiempo	ERP
AS	25049	21.26s	17.7 %
AG	21925	13.69s	3.02 %
memético	21282	1.02s	0.0 %

Tabla III  
PARÁMETROS DEL SISTEMA DE HORMIGAS PARALELO

Parámetro	Valor
Colonias	32
Iteraciones	100
Tamaño de la Colonia	1000
Factor de evaporación	0.1

Tabla IV  
PARÁMETROS DEL AG DISTRIBUIDO

Parámetro	Valor
Subpoblaciones	32
Generaciones	5000
Población	1024
Probabilidad de Selección	50 %
Probabilidad de Mutación	20 %
Elitismo	10
Período Migratorio	50
Tasa de Migración	5

Tabla V  
PARÁMETROS DEL ALGORITMO MEMÉTICO

Parámetro	Valor
Subpoblaciones	32
Generaciones	15
Población	10
Probabilidad de Selección	40 %
Probabilidad de Mutación	1 %
Elitismo	1
Período Migratorio	3
Tasa de Migración	1

Las tablas VI, VII y VIII muestran los resultados obtenidos con el algoritmo propuesto para diferentes instancias del problema tomadas de TSPLIB. El único parámetro que se modifica en cada caso es el tamaño de la población, el cual se muestra en cada tabla.

Tabla VI  
RESULTADOS EXPERIMENTALES INSTANCIAS CHURRITZ Y REINELT (TSP225).

Instancia	Óptimo	Población	ERP	Tiempo
ch130	6110	13	0.0 %	3.2s
ch150	6528	15	0.0 %	3.3s
tsp225	3919	22	0.0 %	21.3s

Tabla VII  
RESULTADOS EXPERIMENTALES INSTANCIAS KROLAK/FELTS/NELSON.

Instancia	Óptimo	Población	ERP	Tiempo
kroA100	21282	10	0.0 %	1.0s
kroB100	22141	10	0.0 %	1.0s
kroC100	20749	10	0.0 %	1.0s
kroD100	21294	10	0.0 %	1.1s
kroE100	22068	10	0.0 %	1.0s
kroA150	26524	15	0.0 %	4.2s
kroB150	26130	15	0.0 %	4.4s
kroA200	29368	20	0.0 %	15.7s
kroB200	29437	20	0.0 %	14.6s

Tabla VIII  
RESULTADOS EXPERIMENTALES INSTANCIAS PADBERG/RINALDI.

Instancia	Óptimo	Población	ERP	Tiempo
pr107	44303	10	0.0 %	3.0s
pr124	59030	12	0.0 %	3.3s
pr136	96772	13	0.0 %	3.1s
pr144	58537	14	0.0 %	9.0s
pr152	73682	15	0.0 %	8.8s
pr226	80369	22	0.0 %	44.3s
pr264	49135	26	0.0 %	94.7s
pr299	48191	30	0.0 %	75.5s

## VIII. CONCLUSIONES Y TRABAJO FUTURO

El algoritmo de búsqueda propuesto parece ser una mejor alternativo al algoritmo 3-opt y a la versión favorable de 2-opt ya que, de acuerdo a la tabla I, la media del costo de la



ruta encontrada por dicho algoritmo es claramente mejor que la obtenida utilizando 2-opt y si bien la media es similar a la encontrada por 3-opt, el algoritmo propuesto es hasta diez veces más rápido y, por tanto más adecuado para llevarse a cabo masivamente en un algoritmo evolutivo poblacional.

El algoritmo memético basado en la estrategia de búsqueda local propuesta encuentra el recorrido óptimo de las instancias de la TSPLIB de hasta 300 ciudades utilizadas como prueba de acuerdo a las tablas VI, VII y VIII. Por otro lado, utilizando un sistema de hormigas paralelo o un algoritmo genético distribuido, incluso para las instancias de 100 ciudades propuestas por Krolak, Felts y Nelson, no se alcanza el óptimo aún en tiempos diez veces mayores que el requerido por el algoritmo memético de acuerdo a la tabla II.

Actualmente se está trabajando en la implementación del algoritmo memético en una arquitectura paralela distinta como son las GPU, que recientemente han obtenido resultados mejores en relación a los *cluster* en algunos algoritmos paralelizables.

Otra cuestión que queda como trabajo futuro es la comparación de la heurística de búsqueda local propuesta contra estrategias más elaboradas presentes en la literatura como la de Lin-Kernighan-Helsgaun, entre otras.

#### REFERENCIAS

- [1] S. Arora, "Polynomial time approximation schemes for Euclidean traveling salesman and other geometric problems," *Journal of the ACM*, vol. 45, no. 5, pp. 753–782, 1998.
- [2] G. Gutin and A. P. Punnen, *The traveling salesman problem and its variations*, vol. 12. Kluwer Academic Publishers, 2002.
- [3] Y. Chen and P. Zhang, "Optimized annealing of traveling salesman problem from the  $n$ th-nearest-neighbor distribution," *Physica A: Statistical Mechanics and its Applications*, vol. 371, no. 2, pp. 627–632, 2006.
- [4] J.-q. Li, Q.-k. Pan, and Y.-C. Liang, "An effective hybrid tabu search algorithm for multi-objective flexible job-shop scheduling problems," *Computers & Industrial Engineering*, vol. 59, no. 4, pp. 647–662, 2010.
- [5] S. J. Louis and G. Li, "Case injected genetic algorithms for traveling salesman problems," *Information sciences*, vol. 122, no. 2, pp. 201–225, 2000.
- [6] K.-S. Leung, H.-D. Jin, and Z.-B. Xu, "An expanding self-organizing neural network for the traveling salesman problem," *Neurocomputing*, vol. 62, pp. 267–292, 2004.
- [7] M. Dorigo and L. M. Gambardella, "Ant colonies for the travelling salesman problem," *Biosystems*, vol. 43, no. 2, pp. 73–81, 1997.
- [8] X. H. Shi, Y. C. Liang, H. P. Lee, C. Lu, and Q. X. Wang, "Particle swarm optimization-based algorithms for TSP and generalized TSP," *Information Processing Letters*, vol. 103, no. 5, pp. 169–176, 2007.
- [9] M. Padberg and G. Rinaldi, "A Branch-and-Cut Algorithm for the Resolution of Large-Scale Symmetric Traveling Salesman Problems," *SIAM Review*, vol. 33, no. 1, pp. 60–100, 1991.
- [10] K. Helsgaun, "An effective implementation of the Lin–Kernighan traveling salesman heuristic," *European Journal of Operational Research*, vol. 126, pp. 106–130, oct 2000.
- [11] G. Clarke and J. W. Wright, "Scheduling of Vehicles from a Central Depot to a Number of Delivery Points," *Operations Research*, vol. 12, no. 4, pp. 568–581, 1964.
- [12] M. Gendreau, A. Hertz, and G. Laporte, "New Insertion and Post Optimization Procedures for the Traveling Salesman Problem," *Operations Research*, vol. 40, no. 6, pp. 1086–1095, 1992.
- [13] Y.-t. Wang, J.-q. Li, K.-z. Gao, and Q.-k. Pan, "Memetic Algorithm based on Improved Inver–over operator and Lin–Kernighan local search for the Euclidean traveling salesman problem," *Computers & Mathematics with Applications*, vol. 62, pp. 2743–2754, oct 2011.
- [14] G. C. Onwubolu, *New Optimization Techniques in Engineering*, ch. Optimizing, pp. 537–565. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004.
- [15] P. Merz, *Memetic Algorithms for Combinatorial Optimization Problems: Fitness Landscapes and Effective Search Strategies*. PhD thesis, 2000.
- [16] P. Moscato and C. Cotta, "A Modern Introduction to Memetic Algorithms," in *Handbook of Metaheuristics*, vol. 146, pp. 141–183, 2010.
- [17] W. Hart, N. Krasnogor, and J. Smith, "Memetic evolutionary algorithms," *Recent advances in memetic algorithms*, pp. 3–27, 2005.
- [18] NVIDIA, *NVIDIA CUDA C Programming Guide*. No. 350, 2014.
- [19] R. L. Graham, T. S. Woodall, and J. M. Squyres, "Open MPI: A Flexible High Performance MPI," *Proceedings 6th Annual International Conference on Parallel Processing and Applied Mathematics*, pp. 228–239, 2005.
- [20] J. A. Castro, M. A. Castro Liera, and I. Castro Liera, *Programación paralela aplicada en optimización*. La Paz, B.C.S.: Instituto Tecnológico de La Paz, 1 ed., 2012.
- [21] M. A. Castro Liera, *Un algoritmo genético distribuido con aplicación en la identificación difusa de un proceso fermentativo*. Ph.d., Universidad Central, "Marta Abreu" de Las Villas, 2009.
- [22] L. Davis, "Job Shop Scheduling with Genetic Algorithms," in *First International Conference on Genetic Algorithms*, (NJ, USA), pp. 136–140, L. Erlbaum Associates Inc., jul 1985.
- [23] G. Reinelt, "TSPLIB—A Traveling Salesman Problem Library," *INFORMS Journal on Computing*, vol. 3, no. 4, pp. 376–384, 1991.
- [24] Y. T. Wang, J. Q. Li, K. Z. Gao, and Q. K. Pan, "Memetic Algorithm based on Improved Inver-over operator and Lin-Kernighan local search for the Euclidean traveling salesman problem," in *Computers and Mathematics with Applications*, vol. 62, pp. 2743–2754, 2011.

# Análisis de desempeño en la implementación de la DCT bidimensional en cómputo paralelo

Jose R. Rosales-Roldan, Rogelio Reyes-Reyes, Clara Cruz-Ramos, Stephanie E. Ramirez-Reyes  
 Maestría en Ingeniería en Seguridad y Tecnologías de la Información  
 Instituto Politécnico Nacional, SEPI ESIME Unidad Culhuacán México, D.F.  
 raosjr3@gmail.com, rreyesre@ipn.mx, ccruzra@ipn.mx, euniceramirez.rs@gmail.com

**Resumen**—La Transformada Discreta de Coseno (DCT por sus siglas en inglés) es una transformada basada en la Transformada Discreta de Fourier pero con la característica de que solamente utiliza números reales, por su propiedad de compactación de información la DCT es popular en áreas como la compresión de videos e imágenes digitales y también es usada durante el proceso de inserción y extracción de marcas de agua; dependiendo del método de implementación de la DCT que se utilice ésta puede representar un costo computacional alto, lo que afecta el tiempo total de procesamiento. Sin embargo, en la última década se ha desarrollado hardware con capacidad de procesamiento en paralelo, siendo los más populares las unidades de procesamiento gráfico (GPU) y los procesadores multinúcleo (CPU), estos pueden ser utilizados para implementar la DCT en paralelo. En este artículo se analiza el desempeño de tres métodos diferentes para el cálculo de la DCT e IDCT bidimensionales, el primero es la DCT tradicional, el segundo está basado en utilizar dos DCT unidimensionales y el último está basado en la aplicación de la matriz DCT. Para cada método se implementan tres versiones: la primera realiza el procesamiento de manera secuencial, la segunda y tercera versión realizan el procesamiento en paralelo mediante procesadores multinúcleo y unidades de procesamiento gráfico (GPUs) respectivamente. Los resultados demuestran que el método de la matriz DCT al ser implementado en paralelo tanto en GPU como en CPU multinúcleo ofrece el menor tiempo de ejecución.

**Keywords**—DCT, IDCT, CPU multinúcleo, GPU, procesamiento de imágenes

## I. INTRODUCCIÓN

Hoy en día es común tomar fotografías, grabar videos en alta resolución y compartirlos a través de Internet; por tal motivo, el procesamiento digital de imágenes se ha convertido en un área de estudio relevante, las imágenes están constituidas por información y por eso surge la necesidad de reducirlas, es decir, usar la menor cantidad de datos para representar la misma imagen. En los últimos años se han realizado numerosas investigaciones relacionadas con las técnicas de compresión de datos, y en algunos métodos propuestos [1-3] se utilizan transformadas como la DCT y DWT para ayudar al proceso de compresión. Por otro lado, muchas técnicas de inserción de marcas de agua visibles [4-6] e invisibles [7-9] hacen uso de este tipo de transformadas para realizar el proceso de inserción y extracción. En general, se hace evidente el uso de la técnica de la DCT en el procesamiento de señales, particularmente para las imágenes digitales se requiere del cálculo de una DCT bidimensional (2D-DCT), dependiendo del método que se utilice para implementarla puede conllevar a un tiempo de procesamiento y costo computacional alto, lo que para algunas aplicaciones no es admisible ya que se

requiere una respuesta en el menor tiempo posible. Tomando en consideración esta problemática se han propuesto diversas soluciones para agilizar el proceso de cálculo de la 2D-DCT, entre las cuales se encuentran la propuesta por Arai [10] que está basada en el cálculo de dos DCT unidimensionales (1D-DCT) y la segunda propuesta de Gilbert S. [11] que utiliza una matriz DCT (M-DCT). Por tal motivo y debido a que actualmente el hardware de cómputo paralelo es accesible y de bajo costo, se propone hacer la implementación de las propuestas anteriores utilizando las arquitecturas en paralelo basadas en GPU y CPU multinúcleo con la finalidad de evaluar el desempeño de estos métodos en las arquitecturas antes mencionadas.

## II. TRANSFORMADA DISCRETA DE COSENO (DCT)

La DCT es una transformada basada en la Transformada Discreta de Fourier, con la particularidad de que solamente utiliza números reales; una de sus principales características es su capacidad de compactar la energía, es decir, la DCT logra acumular la mayor parte de la información en pocos coeficientes transformados [12].

La definición formal de la 1D-DCT de longitud  $N$  está dada por (1).

$$C(u) = \alpha(u) \sum_{x=0}^{N-1} f(x) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \quad (1)$$

La inversa (1D-IDCT) se encuentra definida por (2).

$$f(x) = \sum_{u=0}^{N-1} \alpha(u) C(u) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \quad (2)$$

donde:

$$\alpha(u) = \begin{cases} \sqrt{\frac{1}{N}} & \text{si } u = 0 \\ \sqrt{\frac{2}{N}} & \text{si } u > 0 \end{cases}$$

Para la 2D-DCT de un bloque  $N \times N$  se encuentra definida por (3) [12].

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x, y) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (3)$$

La inversa de la 2D-DCT (2D-IDCT) para un bloque de tamaño  $N \times N$  está definida por (4).

$$f(x, y) = \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \alpha(u)\alpha(v) C(u, v) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (4)$$

### A) DCT Bidimensional

La 2D-DCT es utilizada ampliamente para el procesamiento de imágenes, utilizar bloques de 8x8 para la DCT es común debido a que el estándar de compresión JPEG está basado en la transformada DCT con bloques de 8x8 [13]. A partir de (3) y (4), las ecuaciones (5) y (6) definen matemáticamente el cálculo de la 2D-DCT para un bloque de 8x8.

$$F(u, v) = \frac{1}{4} C(u)C(v) \sum_{x=0}^7 \sum_{y=0}^7 f(x, y) \cos \left[ \frac{(2x+1)u\pi}{16} \right] \cos \left[ \frac{(2y+1)v\pi}{16} \right] \quad (5)$$

$$f(x, y) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) D(u, v) \cos \left[ \frac{(2x+1)u\pi}{16} \right] \cos \left[ \frac{(2y+1)v\pi}{16} \right] \quad (6)$$

donde

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{si } u = 0 \\ 1 & \text{si } u > 0 \end{cases}$$

### B) 2D-DCT utilizando dos veces 1D-DCT.

De acuerdo a Aria [10], la 2D-DCT puede calcularse a partir de la aplicación de la 1D-DCT. Utilizando la propiedad de separabilidad a (3) se obtiene (7).

$$C(u, v) = \alpha(u)\alpha(v) \sum_{x=0}^7 \cos \left( \frac{(2x+1)u\pi}{16} \right) \left[ \sum_{y=0}^7 f(x, y) \cos \left( \frac{(2y+1)v\pi}{16} \right) \right] \quad (7)$$

La ecuación (7) calcula la 2D-DCT utilizando la ecuación (1) inicialmente en las columnas y posteriormente en los renglones, como se puede observar en la figura 1 [14].

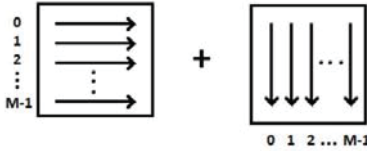


Figura 1: Aplicando 2 veces la 1D-DCT

### C) Matriz DCT

Gilbert S. [11], propone utilizar las ecuaciones (5) y (6) para construir una matriz DCT (M-DCT) para bloques de 8x8, tal y como se muestra en las ecuaciones (8) y (9).

$$T_{i,j} = \begin{cases} \frac{1}{\sqrt{N}} & \text{si } i = 0 \\ \sqrt{\frac{2}{N}} \cos \left[ \frac{(2j+1)i\pi}{2N} \right] & \text{si } i > 0 \end{cases} \quad (8)$$

$$T = \begin{bmatrix} .35 & .35 & .35 & .35 & .35 & .35 & .35 & .35 \\ .49 & .41 & .27 & .09 & -.09 & -.27 & -.41 & -.49 \\ .46 & .19 & -.19 & -.46 & -.46 & -.19 & .19 & .46 \\ .41 & -.09 & -.09 & -.27 & .27 & .49 & .09 & -.41 \\ .35 & -.35 & -.35 & .35 & .35 & -.35 & -.35 & .35 \\ .27 & -.49 & -.49 & .41 & -.41 & -.09 & .49 & -.27 \\ .19 & -.46 & -.46 & -.19 & .19 & .46 & -.46 & .19 \\ .09 & -.27 & .41 & -.49 & .49 & -.41 & .27 & -.09 \end{bmatrix} \quad (9)$$

De acuerdo a Gilbert S. [11], para calcular la DCT y la IDCT empleando la matriz  $T$  se utilizan las ecuaciones (10) y (11) respectivamente.

$$B = TAT' \quad (10)$$

$$A = T'BT \quad (11)$$

## III. PROGRAMACIÓN EN PARALELO

Tradicionalmente el software se desarrolla para ser implementado en computadoras con un solo procesador, las instrucciones se ejecutan una tras otra y solo una instrucción puede ser ejecutada en un momento determinado, es decir, la programación es secuencial. La programación en paralelo o paralelismo es el uso simultáneo de múltiples unidades de procesamiento para resolver un problema, la idea básicamente es dividir un problema grande en varios pequeños que son procesados concurrentemente, es decir, las instrucciones de cada uno de los problemas se ejecutan en diferentes unidades de procesamiento al mismo tiempo [15].

### A) OpenMP

OpenMP es una API que permite añadir concurrencia a las aplicaciones que se ejecutan en multiprocesadores de memoria compartida. Se basa en la creación de hilos de ejecución paralelos compartiendo las variables del proceso principal que los crea. OpenMP ofrece una API de memoria compartida basada en directivas. En C y C++, esto significa que hay instrucciones de preprocesador especiales conocidos como pragmas. Los pragmas se añaden a un sistema para permitir comportamientos que no son parte de la especificación básica [15].

### B) CUDA

CUDA es un modelo de plataforma de computo y programación en paralelo de propósito general que aprovecha el motor de cálculo paralelo de las GPU NVIDIA para resolver muchos problemas computacionales complejos de una manera más eficiente. Utilizando CUDA se puede acceder a la GPU para realizar cálculos numéricos, como se hace tradicionalmente en la CPU. CUDA C es una extensión del estándar C con un conjunto de extensiones de lenguaje para permitir la programación en unidades de procesamiento gráfico, también provee APIs sencillas para administrar dispositivos, memoria y otras tareas. Para aprovechar el computo en paralelo que ofrecen las tarjetas gráficas, CUDA aporta un modelo de programación que se basa en hilos agrupados en bloques que se ejecutan en paralelo para realizar un conjunto de tareas [16].

### C) Factor de aceleración (speedup)

El speedup es una métrica de rendimiento que se puede definir como un indicador que muestra cuantas veces un programa en paralelo es más rápido que su correspondiente programa secuencial. El speedup toma como referencia el tiempo de ejecución de un programa en un sistema secuencial respecto a la medida de tiempo de ejecución del programa en paralelo, esto se encuentra expresado en la ecuación (12) de acuerdo a [14].

$$sp = \frac{T_i}{T_p} \quad (12)$$

donde  $T_i$  representa el tiempo que tarda en ejecutarse el programa para un equipo con un solo procesador y  $T_p$  es el tiempo que tarda en ejecutarse el mismo programa en paralelo.

## IV. MÉTODOS EVALUADOS

Para las pruebas que se realizaron se consideraron tres métodos para calcular la DCT y su correspondiente inversa.

- Bidimensional (2D-DCT)
- Aplicando 2 veces la unidimensional (2 1D-DCT)
- Utilizando la matriz DCT (M-DCT)

Para cada uno de los métodos consideradores se propusieron tres implementaciones: Secuencial, multinúcleo y GPU.

## A) 2D-DCT

El primer método está basado en las ecuaciones (5) y (6) para calcular la DCT e IDCT de un bloque de  $8 \times 8$ .

1) *Secuencial*: La implementación secuencial se muestra en el algoritmo 1.

**Algoritmo 1** 2D-DCT y 2D-IDCT secuencial

---

```

procedure DCT2D(BLOI, U,V)
  res = 0
  for x=0: 7 do
    for y=0: 7 do
      res = res+(bloi[x,y]*cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25*C(u)*C(v))

procedure IDCT2D(BLOI, X,Y)
  res = 0
  for u=0: 7 do
    for v=0: 7 do
      res = res+(C(u)*C(v)*bloi[u,v]*cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25)

```

---

2) *Multinúcleo*: El proceso de cálculo de la 2D-DCT y 2D-IDCT se realiza en cada una de las  $(u, v)$  y  $(x, y)$  iteraciones que son independientes de las demás, por lo que conviene utilizar un ciclo *for* paralelo. Utilizando la directiva *parallel for* de OpenMP, se divide el número de iteraciones entre el número de hilos disponibles, además, debido a que se tiene un *for* anidado es importante utilizar la cláusula *collapse* para aumentar el número de iteraciones que se particionaran entre los hilos. Por último, es importante especificar que recursos serán compartidos por todos los hilos y aquellos que se utilizan de manera privada. El algoritmo 2 muestra la implementación en paralelo sobre CPU multinúcleo del algoritmo 1.

**Algoritmo 2** 2D-DCT y 2D-IDCT multinúcleo

---

```

procedure DCT2DC(BLOI, U,V)
  res = 0
  #pragma omp parallel for shared(u,v,bloi) private(x,y,res) collapse(2)
  for x=0: 7 do
    for y=0: 7 do
      res = res+(bloi[x,y]*cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25*C(u)*C(v))

procedure IDCT2DC(BLOI, X,Y)
  res = 0
  #pragma omp parallel for shared(x,y,bloi) private(u,v,res) collapse(2)
  for u=0: 7 do
    for v=0: 7 do
      res=res+(C(u)*C(v)*bloi[u,v] *cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25)

```

---

3) *GPU*: Para la implementación en paralelo utilizando CUDA se consideró a cada uno de los bloques de la DCT e IDCT como un bloque de hilos, cada hilo del bloque calcula

el coeficiente de cada uno de pixeles que le corresponde. Además, para reducir la lectura en memoria se considera el uso de memoria compartida para minimizar el acceso a memoria global. El algoritmo 3 ilustra la implementación en paralelo sobre GPU del algoritmo 1.

**Algoritmo 3** 2D-DCT y 2D-IDCT GPU

---

```

threads(8,8)
bloks((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure DCT2DG(BLOI, U,V)
  res = 0
  for x=0: 7 do
    for y=0: 7 do
      res=res+(bloi[x,y]*cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25*C(u)*C(v))

threads(8,8)
bloks((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure IDCT2DG(BLOI, X,Y)
  res = 0
  for u=0: 7 do
    for v=0: 7 do
      res=res+(C(u)*C(v)*bloi[u,v] *cos(((2x+1)PIu)/16)*cos(((2y+1)PIv)/16))
  return (res * 0.25)

```

---

## B) 2 1D-DCT

El segundo método está basado en las ecuaciones (1), (2) y (7).

1) *Secuencial*: El algoritmo 4 corresponde a la implementación secuencial de la DCT e IDCT aplicando dos veces la 1D-DCT y 1D-IDCT respectivamente.

**Algoritmo 4** 1D-DCT y 1D-IDCT secuencial

---

```

procedure DCT21D(BLOI, U,V)
  res = 0
  for a=0: 7 do
    res = res + bloi[a+8*v] * cos(((2*a+1)*u*PI)/16)
    matr[u+8*v] = round(res*C(u)*0.5)
  res =0.0
  for b=0: 7 do
    res = res + matr[u+8*b] * cos(((2*b+1)*v*PI)/16)
  return (res*C(v)*0.5)

procedure IDCT21D(BLOI, X,Y)
  res = 0
  for a=0: 7 do
    res = res + Cc(a)*bloi[a+8*y] * cos((2*x+1)*a*PI/16)
    matr[x+8*y] = round(res*.5)
  res =0.0
  for b=0: 7 do
    res = res + Cc(b)*matr[x+8*b] * cos((2*y+1)*b*PI/16)
  return (res*0.5)

```

---

2) *Multinúcleo*: Como se puede observar en el algoritmo 4, cada uno de las iteraciones de  $a$  y  $b$  son independientes, además, al realizar dos veces la 1D-DCT y 1D-IDCT se elimina el *for* anidado, por lo que en este caso es ideal utilizar la directiva *parallel for* para dividir el número de iteraciones de  $a$  y  $b$  entre el número de hilos disponibles. El algoritmo 5 muestra la implementación en paralelo sobre CPU multinúcleo del algoritmo 4.

**Algoritmo 5** 1D-DCT y 1D-IDCT multinúcleo

```

procedure DCT21DC(BLOI, U, V)
  res = 0
  #pragma omp parallel for shared(u,v,bloi) private(a,res)
  for a=0: 7 do
    res = res + bloi[a+8*v] * cos(((2*a+1)*u*PI)/16)
  matr[u+8*v] = round(res*C(u)*0.5)
  res = 0.0
  #pragma omp parallel for shared(u,v,bloi) private(b,res)
  for b=0: 7 do
    res = res + matr[u+8*b] * cos(((2*b+1)*v*PI)/16)
  return (res*C(v)*0.5)

procedure IDCT21DC(BLOI, X, Y)
  res = 0
  #pragma omp parallel for shared(x,y,bloi) private(a,res)
  for a=0: 7 do
    res = res + Cc(a)*bloi[a+8*y] * cos((2*x+1)*a*PI/16)
  matr[x+8*y] = round(res*.5)
  res = 0.0
  #pragma omp parallel for shared(x,y,bloi) private(b,res)
  for b=0: 7 do
    res = res + Cc(b)*matr[x+8*b] * cos((2*y+1)*b*PI/16)
  return (res*0.5)

```

3) *GPU*: Para la implementación del algoritmo 4 en GPU se considera cada bloque DCT como un bloque de hilos, cada hilo del bloque tiene acceso a un espacio de memoria compartida, además se sincronizan para actualizar los resultados del cálculo de las columnas para posteriormente calcular el resultado de los renglones. El algoritmo 6 muestra la implementación de 4 en la GPU.

**Algoritmo 6** 1D-DCT y 1D-IDCT GPU

```

threads(8,8)
blocs((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure DCT21DG(BLOI, U, V)
  res = 0
  for a=0: 7 do
    res = res + bloi[a+8*v] * cos(((2*a+1)*u*PI)/16)
  matr[u+8*v] = round(res*C(u)*0.5)
  __synsc()
  res = 0.0
  for b=0: 7 do
    res = res + matr[u+8*b] * cos(((2*b+1)*v*PI)/16)
  return (res*C(v)*0.5)

threads(8,8)
blocs((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure IDCT21DG(BLOI, X, Y)
  res = 0
  for a=0: 7 do
    res = res + Cc(a)*bloi[a+8*y] * cos((2*x+1)*a*PI/16)
  matr[x+8*y] = round(res*.5)
  __synsc()
  res = 0.0
  for b=0: 7 do
    res = res + Cc(b)*matr[x+8*b] * cos((2*y+1)*b*PI/16)
  return (res*0.5)

```

**C) M-DCT**

El tercer y último método está basado en las ecuaciones (10) y (11) y utiliza la matriz DCT que se obtuvo en (9) a partir de (8).

1) *Secuencial*: El algoritmo 7 ilustran la implementación secuencial del método propuesto.

**Algoritmo 7** M-DCT secuencial

```

procedure DCTMD(BLOI, U, V, MATRIXD, MATRIXID)
  res = 0
  for c=0: 7 do
    res = res + (MatrixD[u*8+c]* bloi[c*8+v])
  matr[u+8*v] = res
  res = 0.0
  for c=0: 7 do
    res = res + (matr[u*8+c] *MatrixID[c*8+v])
  return res

procedure IDCTMD(BLOI, X, Y, MATRIXD, MATRIXID)
  res = 0
  for c=0: 7 do
    res = res + (MatrixID[x*8+c]* bloi[c*8+y])
  matr[x+8*y] = res
  res = 0.0
  for c=0: 7 do
    res = res + (matr[x*8+c] *MatrixD[c*8+y])
  return res

```

2) *Multinúcleo*: Como se observa en el algoritmo 7, cada una de las iteraciones de  $a$  y  $b$  se realiza de manera independiente, por que el **for** puede ser paralelizado utilizando la directiva **parallel for**. El algoritmo 8 muestra la implementación en paralelo sobre CPU multinúcleo del algoritmo 7.

**Algoritmo 8** M-DCT multinúcleo

```

threads(8,8)
blocs((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure DCTMDC(BLOI, U, V, MATRIXD, MATRIXID)
  res = 0
  #pragma omp parallel for shared(u,v,bloi) private(c,res)
  for c=0: 7 do
    res = res + (MatrixD[u*8+c]* bloi[c*8+v])
  matr[u+8*v] = res
  res = 0.0
  #pragma omp parallel for shared(u,v,bloi) private(c,res)
  for c=0: 7 do
    res = res + (matr[u*8+c] *MatrixID[c*8+v])
  return res

threads(8,8)
blocs((cols+threads.x-1)/threads.x,(rows+threads.y-1)/threads.y)
Hacer en paralelo
procedure IDCTMDC(BLOI, X, Y, MATRIXD, MATRIXID)
  res = 0
  #pragma omp parallel for shared(x,y,bloi) private(c,res)
  for c=0: 7 do
    res = res + (MatrixID[x*8+c]* bloi[c*8+y])
  matr[x+8*y] = res
  res = 0.0
  #pragma omp parallel for shared(x,y,bloi) private(c,res)
  for c=0: 7 do
    res = res + (matr[x*8+c] *MatrixD[c*8+y])
  return res

```

3) *GPU*: Para la implementación del algoritmo 7 en GPU se considera cada bloque DCT como un bloque de hilos, cada hilo del bloque tiene acceso a un espacio de memoria compartida, además se sincronizan para actualizar los resultados del cálculo de la matriz DCT para posteriormente calcular el resultado de la multiplicación de la matriz DCT inversa. El algoritmo 9 muestra la implementación de 7 en la GPU.

**Algoritmo 9** M-DCT GPU

```

1: procedure DCTMDG(BLOI, U,V, MATRIXD, MATRIXID)
2:   res = 0
3:   for c=0: 7 do
4:     res = res + (MatrixD[u*8+c]* bloi[c*8+v])
5:     __synsc()
6:   matr[u+8*v] = res
7:   res =0.0
8:   for c=0: 7 do
9:     res = res + (matr[u*8+c] *MatrixID[c*8+v])
10:    __synsc()
11:   return res
1: procedure IDCTMDG(BLOI, X,Y,MATRIXD, MATRIXID)
2:   res = 0
3:   for c=0: 7 do
4:     res = res + (MatrixID[x*8+c]* bloi[c*8+y])
5:     __synsc()
6:   matr[x+8*y] = res
7:   res =0.0
8:   for c=0: 7 do
9:     res = res + (matr[x*8+c] *MatrixD[c*8+y])
10:    __synsc()
11:   return res
    
```

V. RESULTADOS

Para medir el rendimiento de los algoritmos propuestos se utilizó el factor de aceleración que está definido por la ecuación (12). A continuación, se presentan los resultados obtenidos de los algoritmos propuestos implementados en un equipo de cómputo con las características que se listan en la tabla I.

Tabla I: Características del equipo utilizado

Elemento	Descripción
CPU	Intel Core i7
GPU	NVIDIA GeForce GTX 980 Ti
RAM	16 Gb DDR4 2200 Mhz
Almacenamiento	SSD Kingston HyperX Fury 480 Gb
Sistema operativo	Linux Debian Kernel
Compilador	GNU g++ 4.9. NVIDIA CUDA nvcc 7.5

Para las pruebas realizadas, se utilizaron imágenes de NxN (512x512, 1024x1024, 2048x2048, 4096x4096, 6144x6144, 8192x8192 y 10240x10240). Se les aplico la DCT y su inversa, para cada implementación se diseñaron tres propuestas: secuencial y en paralelo en la CPU utilizando OpenMP y en paralelo en GPU utilizando CUDA. Los algoritmos se ejecutaron 100 veces y se obtuvo el promedio. A continuación, en las tablas II y III se muestran los tiempos de ejecución para el cálculo de la DCT e IDCT bidimensional.

Tabla II: Tiempos de ejecución en segundos para la 2D-DCT

Tipo	512	1024	2048	4096	6144	8192	10240
Secuencial	0.9730	3.1877	12.992	52.983	121.38	226.15	304.43
Paralelo	0.1929	0.6037	2.3257	8.9360	19.502	34.072	44.330
GPU	0.0004	0.0017	0.0066	0.0217	0.0489	0.1045	0.1550

Tabla III: Tiempos de ejecución en segundos para la 2D-IDCT

Tipo	512	1024	2048	4096	6144	8192	10240
Secuencial	0.9731	3.1884	12.993	52.991	121.39	227.00	304.51
Paralelo	0.21660	0.6212	2.5426	8.9654	19.642	34.562	44.458
GPU	0.0004	0.0016	0.0066	0.0221	0.0499	0.0959	0.1389

Los resultados que se muestran en las tablas II y III confirman que la implementación en la CPU secuencial es la más

lenta, mientras que la implementación en CPU multinúcleo muestra que es entre 5.0 y 6.8 veces más rápida que su contraparte secuencial, por otro lado, la GPU ofrece una mejor aceleración con respecto a la versión secuencial y en paralelo, ya que su factor de aceleración es entre 1900 y 2400 veces más rápida que la propuesta secuencial y con respecto a la implementación en CPU multinúcleo es entre 280 y 430 veces más rápida.

La figura 2 muestra el factor de aceleración para la 1D-DCT y 1D-IDCT implementadas de manera secuencial, en paralelo y en GPU.

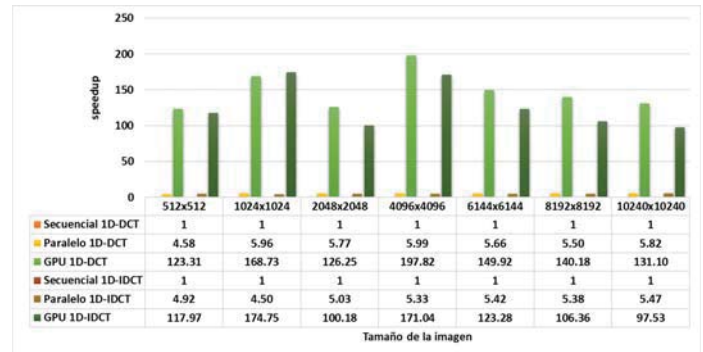


Figura 2: Factor de aceleración para la 1D-DCT y 1D-IDCT en secuencial, paralelo y GPU

Como se observa en la figura 2, la 1D-DCT y 1D-IDCT implementada en paralelo es entre 4.58 y 5.99 veces más rápida que la versión secuencial, mientras que la implementación en GPU es 97.53 y 197.82 veces más rápida que la versión secuencial.

La figura 3 muestra el factor de aceleración para la M-DCT y M-IDCT implementadas de manera secuencial, en paralelo y en GPU.

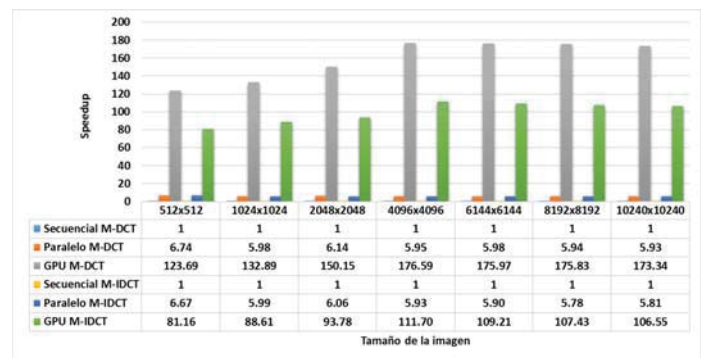


Figura 3: Factor de aceleración para la M-DCT y M-IDCT en secuencial, paralelo y GPU

Se puede observar en la figura 3 que el factor de aceleración para la versión en paralelo es de entre 5.98 y 6.74 respecto a la versión secuencial, mientras que la implementación en GPU es entre 81.16 y 175.97 veces más rápida que la implementación secuencial.

Como se observa en las tablas II, III y en las figuras 2, 3 la 2D-DCT, 1D-DCT y M-DCT con sus respectivas inversas implementadas en la GPU ofrecen el mejor rendimiento en comparación con las secuenciales y en paralelo. Debido a los resultados de las tablas II y III, la propuesta de la 2D-DCT con su inversa son descartadas como el método más rápido. Ahora bien, la figura 4 ilustra los mejores tiempos de ejecución de la 1D-DCT y M-DCT con sus respectivas inversas implementadas en CPU.

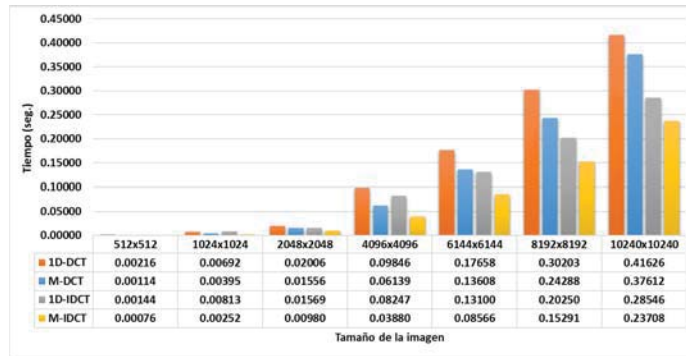


Figura 4: Tiempos de ejecución para la 1D-DCT, 1D-IDCT, M-DCT y M-IDCT en CPU

Como se puede observar en la figura 4, la M-DCT y su inversa son las más rápidas ya que muestran el menor tiempo de ejecución y son entre 1.2 y 3.22 veces más rápidas que la 1D-DCT. La figura 5 muestra los tiempos de ejecución de la 2D-DCT, 1D-DCT y M-DCT con sus respectivas inversas implementadas en la GPU.

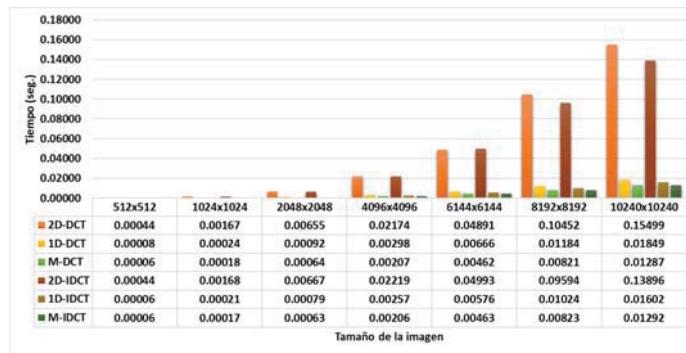


Figura 5: Tiempos de ejecución para la 2D-DCT, 2D-IDCT, 1D-DCT, 1D-IDCT, M-DCT y M-IDCT en GPU

Los resultados, al igual que en la versión en paralelo, la M-DCT con su respectiva inversa son más rápidas que la 2D-DCT y la 1D-DCT, en promedio es entre 7.32 y 9.36 veces más rápida que la 2D-DCT y entre .96 y 1.25 veces más rápida que la 1D-DCT. Además, los resultados de las figuras 4 y 5 muestran que la implementación en GPU con el peor rendimiento es entre 1.46 y 1.74 veces más rápida que la mejor implementación en CPU multinúcleo.

## VI. CONCLUSIONES

En este artículo se presentaron tres métodos para calcular la DCT e IDCT a un bloque de 8x8, los cuales presenta-

ron tres implementaciones diferentes, la primera realiza el procesamiento secuencial, la segunda utiliza un procesador multinúcleo para realizar el cálculo en paralelo y, por último, la tercera versión realiza el procesamiento en paralelo utilizando una unidad de procesamiento gráfico (GPU). En base a los resultados, se puede concluir que el mejor método para calcular la DCT e IDCT es utilizando la Matriz DCT implementada en la GPU ya que resulta ser entre 12.16 y 18.85 veces más rápida que la mejor versión implementada en la CPU multinúcleo. Los resultados también muestran que la DCT bidimensional es el método más ineficiente para el cálculo de la DCT ya que es entre 806.07 y 963.08 veces más lento que la mejor versión en CPU multinúcleo y 15601.15 y 27544.92 veces más lento que la mejor versión en GPU. Además, la peor implementación en GPU es entre 1.46 y 1.74 veces más rápida que la mejor implementación en CPU multinúcleo. Por último, cualquiera de los métodos implementados en la GPU es a lo menos 12.16 veces más rápido que la mejor implementación en CPU.

## AGRADECIMIENTOS

Los autores agradecen al Instituto Politécnico Nacional, a la COFAA, y al CONACYT por el apoyo ofrecido para la realización de esta investigación.

## REFERENCIAS

- [1] Shapiro, J.M, *Embedded Image Coding Using Zerotrees of Wavelet Coefficients* (1993) IEEE Transactions on Signal Processing, 41 (12), pp. 3445-3462
- [2] Zhang, X., Cheng, H. *Histogram-based retrieval for encrypted JPEG images* (2014) 2014 IEEE China Summit and International Conference on Signal and Information, IEEE ChinaSIP 2014 - Proceedings, art. no. 6889282, pp. 446-449.
- [3] Cheng, H., Zhang, X., Yu, J., Li, F. *Markov process based retrieval for encrypted JPEG images* 0.4em(2015) Availability, Reliability and Security (ARES), 2015 10th International Conference on, pp. 417-421. IEEE, Toulouse
- [4] Chen, C., Shi, Y.Q. *JPEG image steganalysis utilizing both intrablock and interblock correlations* (2008) Proceedings - IEEE International Symposium on Circuits and Systems, art. no. 4542096, pp. 3029-3032.
- [5] Chu, W.C. *DCT-based image watermarking using subsampling* (2003) IEEE Transactions on Multimedia, 5 (1), pp. 34-38. Cited 170 times. doi: 10.1109/TMM.2003.808816
- [6] Kang, X., Huang, J., Shi, Y.Q., Lin, Y. *A DWT-DFT Composite Watermarking Scheme Robust to Both Affine Transform and JPEG Compression* (2003) IEEE Transactions on Circuits and Systems for Video Technology, 13 (8), pp. 776-786.
- [7] Singh, A.K., Dave, M., Mohan, A. *Hybrid technique for robust and imperceptible image watermarking in DWT-DCT-SVD domain* (2014) National Academy Science Letters, 37 (4), pp. 351-358.
- [8] Strycker, L., Termont, P., Vandeweghe, J., Haitsma, J., Kalker, A., Maes, M., Depovere, G. *Implementation of a real-time digital watermarking process for broadcast monitoring on a TriMedia VLIW processor* (2000) IEE Proceedings: Vision, Image and Signal Processing, 147 (4), pp. 371-376.
- [9] Salman, A.G., Kanigoro, B., Rojali, Kevin, Santy *Application hiding messages in JPEG images with the method of bit-plane complexity segmentation on android-based mobile devices* (2012) Procedia Engineering, 50, pp. 314-324
- [10] Y. Arai, T. Agui, M. Nakajima. *A fast DCT-SQ scheme form images* Transactions of IECE, Vol. E71, n11, pp.1095-1097, 1998
- [11] Gilbert S. *The Discrete Cosine Transform* SIAM Review, of Society for Industrial and Applied Mathematics Philadelphia, USA Volume 41, Pages 135-147, 1999.
- [12] Nagaria, B., Hashmi, M.F., Patidar, V., Jain, N. *An optimized fast discrete cosine transform approach with various iterations and optimum numerical factors for image quality evaluation* 2011 International Conference on Computational Intelligence and Communication Systems, art. no. 6112846, pp. 158-162.
- [13] Diego Santa Cruz, Mathias Larsson. *JPEG2000, the Next Millenium Compression Standard for Still Images*. Proceedings of the IEEE International Conference on Multimedia Computing and Systems, Maryline Charrier.
- [14] John G. Proakis, Dimitris G. Monolakis, *Digital Signal Processing Principles, Algorithms, and Applications (Third Edition)*
- [15] Peter S. Pacheco. *An Introduction to Parallel Programming* Morgan Kaufmann Publishers is an imprint of Elsevier. 2011.
- [16] John Cheng, Max Grossman and Ty McKercher *Professional CUDA C Programming* Published by John Wiley and Sons, Inc. 2014.

# *Algoritmo de agrupamiento aplicado en la inteligencia de negocios usando GPGPU.*

ISC. Misael Cosio Dominguez  
Instituto Tecnológico de La Paz  
ITLP  
La Paz, Baja California Sur, México  
M14310016@itlp.edu.mx

MSC. Iliana Castro Liera  
Instituto Tecnológico de La Paz  
ITLP  
La Paz, Baja California Sur, México  
icastro@itlp.edu.mx

MATI. Luis Armando Cardenas Florido  
Instituto Tecnológico de La Paz  
ITLP  
La Paz, Baja California Sur, México  
Armando.cardenas@itlp.edu.mx

MC. Jesús Antonio Castro  
Instituto Tecnológico de La Paz  
ITLP  
La Paz, Baja California Sur, México  
jcastro@itlp.edu.mx

DR. Marco Antonio Castro Liera  
Instituto Tecnológico de La Paz  
ITLP  
La Paz, Baja California Sur, México  
mcastro@itlp.edu.mx

*Resumen* – Se plantea una técnica de programación paralela (GPGPU) para la optimización del tiempo de ejecución de un algoritmo de agrupamiento en el procesamiento de información aplicado a la minería de datos, con el objetivo de encontrar un comportamiento de la información de una tienda de autoservicio.

*Palabras Clave* — *Inteligencia de negocios, minería de datos, algoritmos de agrupamiento, K-means, GPU, programación paralela.*

## I. INTRODUCCIÓN

El procesamiento de grandes volúmenes de datos en un ambiente comercial y la interpretación de la información resultante requiere la utilización de herramientas computacionales de alto desempeño. La minería de datos aplicada a la inteligencia de los negocios [1] es una herramienta clave para encontrar factores importantes que permitan establecer el comportamiento de la información a través del tiempo y, a su vez, determinar un patrón entre los datos, con el fin de realizar una toma de decisiones de manera eficiente.

Los sistemas de información trabajan frecuentemente con grandes volúmenes de datos, y explotar ese volumen de información para llevar a cabo una toma de decisiones en un periodo de tiempo considerablemente corto tiene costos muy elevados para las organizaciones que utilizan hardware de capacidad limitada. El incremento del tiempo de procesamiento de la información se convierte en una gran limitante para obtener un mejor desempeño [2].

Una manera de mejorar el desempeño en el tiempo de procesamiento es la utilización de la técnica de procesamiento paralelo GPGPU (General Purpose Graphics Processing Unit).

Este proyecto se llevó a cabo en la División de Estudios de Posgrado e Investigación en el Instituto Tecnológico de La Paz. Los datos fueron proporcionados por la tienda de autoservicio Castores.

## II. OBJETIVO

Paralelizar un algoritmo de agrupamiento utilizando técnicas de programación paralela con CUDA para comparar su eficiencia y eficacia contra resultados obtenidos tradicionalmente.

## III. MATERIALES Y MÉTODOS

Para esta investigación se utilizó un algoritmo de minería de datos aplicado en la inteligencia de negocios, en combinación con las nuevas técnicas de programación paralela en una arquitectura CUDA.

El desarrollo de este proyecto se dividió en 3 secciones: la primera sección involucra los datos, en la segunda sección se trabajó el algoritmo de agrupamiento y, finalmente, el procesamiento paralelo.

### A. Datos para la investigación

Para este proyecto se seleccionó una muestra experimental de 1,000,000 de registros de una base de datos de 19 millones de registros en SQL SERVER 2012. Cada registro posee 15 características del producto y a cada registro se le aplicó el proceso ETL (Extracción, Transformación y Carga) utilizado en la minería de datos como una herramienta que tiene como finalidad mantener una limpieza homogénea y coherencia entre los datos, para minimizar el margen error en su procesamiento.

Cabe mencionar que los registros seleccionados solamente comprenden las transacciones realizadas en un lapso de tiempo de 12 meses, en el Departamento de Licores del negocio.



Cuando se trabaja con grandes cantidades de información, comprender los datos es un factor importante para el éxito en una toma de decisiones. Se creó un cubo multidimensional utilizando el complemento SQL DATA TOOL 2012 en VISUAL STUDIO 2012, para analizar de forma detallada los comportamientos que tienen los datos entre sí y, además, visualizar las características que poseen cada uno de esos comportamientos. Los comportamientos se determinan a partir del funcionamiento del algoritmo de agrupamiento implementado en la minería de datos para la inteligencia de negocios.

*B. Algoritmo de agrupamiento*

Existe una gran variedad de algoritmos para la minería de datos o procesamiento de información para la inteligencia de negocios. En la tabla 1 se encuentran clasificados los algoritmos más utilizados en la minería de datos.

Nombre de algoritmos	PREDICTIVO		DESCRIPTIVO		
	Clasificación	Regresión	Agrupamiento	Reglas de asociación	Combinaciones / Factorizaciones
Redes neuronales	✓	✓	✓		
Arboles de decisión ID3, C4.5, C5.0	✓				
Arboles de decisión CART	✓	✓			
Otros árboles de decisión	✓	✓	✓	✓	
Redes de kohonen			✓		
Regresión lineal y logarítmica		✓			✓
Regresión logística	✓			✓	
K-means			✓		
A priori				✓	
Naive Bayes	✓				
Vecinos más próximos	✓	✓	✓		
Análisis factorial y de comp. ppales					✓
Twostep y Cobweb			✓		
Algoritmos genéticos y evolutivos	✓	✓	✓	✓	✓
Máquinas de vectores soporte	✓	✓	✓		
CN2 rules (cobertura)	✓			✓	
Análisis discriminante multivariante	✓				

Tabla 1.- Algoritmos implementados en minería de datos

Cada uno de estos algoritmos funciona de forma diferente, por lo que son tratados bajo condiciones controladas para problemas con un objetivo específico. Es difícil determinar o evaluar este tipo de algoritmos porque el desempeño que muestre un algoritmo depende de la problemática a resolver.

Para el desarrollo de esta investigación se trabajó con una técnica de agrupamiento (conocida en la literatura como clustering) cuyo objetivo es obtener grupos o conjuntos entre elementos de un universo de datos, de tal manera que los elementos asignados al mismo conjunto posean características similares.

En la tabla 1 se muestran los algoritmos descriptivos que trabajan con la técnica de agrupamiento. Dentro de esa clasificación se encuentra el algoritmo K-means. Este algoritmo es clasificado en la literatura como un algoritmo fácil de paralelizar [1].

K-means es un método de agrupamiento por vecindad en el que se parte de un número determinado de conjuntos y un determinado número de elementos por agrupar.

El objetivo de K-means es situar los K-centroides en el espacio, de forma que los datos más cercanos a determinado K-centroides formen una agrupación de datos a partir de una característica de similitud determinada por una función de distancia [1].

El desempeño del algoritmo K-means puede variar por la función distancia para un mismo problema. Una función distancia puede obtener mejores resultados que otra función distancia que se implemente al mismo algoritmo K-means, para el mismo problema a resolver.

Existen diferentes funciones de distancia que se pueden adaptar al algoritmo K-means; las más comúnmente usadas son:

- Distancia Euclídea

$$d(x, y) = \sqrt{\sum_{i=1}^n (X_i - y_i)^2}$$

- Distancia de Manhattan

$$d(x, y) = \sum_{i=1}^n |X_i - y_i|$$

- Distancia de Chebychev

$$d(x, y) = \max_{i=1 \dots n} |X_i - y_i|$$

- Distancia del Coseno

$$d(x, y) = \arccos \left( \frac{X^T y}{(\|X\| \cdot \|y\|)} \right)$$

- Distancia de Mahalanobis

$$d(x, y) = \sqrt{(x - y)^T S^{-1} (x - y)}$$

Estas funciones son las más conocidas, por ser de naturaleza sencilla y que trabajan con dos instancias X, Y. En un problema estas instancias representan dos características de un dato. Para la solución a nuestro problema se decidió implementar la función distancia euclídea, porque es la función más conocida y más sencilla de comprender. Además, la función distancia euclídea, por su definición, se basa en el concepto más común de distancia usado en el área de las matemáticas.

### C. Procesamiento paralelo

El procesamiento paralelo surgió de la necesidad de ejecutar múltiples procesos que permitan realizar una gran cantidad de operaciones en un periodo de tiempo muy pequeño.

CUDA es una plataforma de procesamiento paralelo y un modelo de programación creado por la compañía NVIDIA, que permite aprovechar al máximo el nivel de procesamiento que tienen las GPU NVIDIA [4].

Si una GPU tiene la capacidad de ejecutar un gran número de operaciones en un segundo, podemos realizar millones de operaciones simultáneamente en un grupo de GPGPU trabajando al mismo tiempo. Es por esto que al procesar grandes volúmenes de información utilizando las técnicas de procesamiento paralelo bajo la arquitectura CUDA se puede reducir de manera considerable el tiempo de procesamiento.

Para reducir el tiempo de procesamiento en esta investigación se desarrolló e implementó el algoritmo K-means en la arquitectura CUDA.

Se creó una estructura en la CPU que funciona como un arreglo para almacenar la información que se leyó del archivo fuente de texto plano. Una vez terminada la operación de lectura, se copió la información a una variable que se utilizó por la GPU. Posteriormente, se eligieron las características más relevantes y se aplicó una normalización a cada dato de la muestra para mantener los valores en rangos entre 0 y 1, para que los datos no presentaran valores muy altos y evitar un desplazamiento mayor de los K-centroides. Finalmente, se realizó el cálculo de la función euclídea con cada uno de los registros, utilizando cada uno de los K-centroides determinados inicialmente de forma secuencial.

## IV. RESULTADOS Y CONCLUSIONES

Solo se ha considerado una prueba experimental para conocer el comportamiento de K-means en nuestro universo de datos en el Departamento de Licores.

Con el análisis detallado de SQL DATA TOOL, se obtuvieron los resultados mostrados.

Departamento Licores	
Total de registros	1,374,381
N° Clúster iniciales	10
% de similitud	100
Clúster con relación fuerte	clúster 2 y clúster 5

Tabla 2.- Información para el agrupamiento a 100% de similitud

En la tabla 2 se muestra la configuración del agrupamiento a 100 % de similitud, obteniendo una relación entre el clúster 2 y el clúster 5.

CLUSTER 2			CLUSTER 5		
Precio venta	0-45,7	99,781 %	Precio venta	0-45,7	100 %
Ventas	0-147,8	93,588 %	Ventas	0-147,8	100 %
Familia producto	JUGO/NECTAR	80,154 %	Familia producto	JUGO/NECTAR	72,297 %
Fecha compra	21-02-2013 al 27-09-2013	25,744 %	Fecha compra	21-02-2013 al 27-09-2013	25,778 %
	27-09-2013 al 03-05-2014	25,179 %		27-09-2013 al 03-05-2014	25,642 %
	03-05-2014 al 06-03-2015	18,955 %		03-05-2014 al 06-03-2015	19,429 %

Tabla 3.- Características de la relación en clúster 2 y clúster 5

La relación que tienen el clúster 2 y el clúster 5 está determinada por los valores que se presentan en las características de precio venta, ventas y familia del producto, como se muestra en la tabla 3.

Departamento Licores	
Total de registros	1,374,381
N° Clúster iniciales	10
% de similitud	80
Clúster con relación fuerte	clúster 2 y clúster 5 clúster 6 y clúster 7

Tabla 4.- Información para el agrupamiento a 80% de similitud

Aplicando una configuración a 80 % de similitud entre los agrupamientos, como se muestra en la tabla 4, se encontró otro agrupamiento de la relación, el cual se muestra la tabla 5.

CLUSTER 6			CLUSTER 7		
Precio venta	45,7 - 131,8	73,127 %	Precio venta	131,8 - 428,5	59,856 %
Ventas	0-147,8	91,070 %	Ventas	147,8 - 608,8	75,164 %
Familia producto	VINO	33,983 %	Familia producto	VINO	28,272 %
Fecha compra	21-02-2013 al 27-09-2013	24,397 %	Fecha compra	21-02-2013 al 27-09-2013	22,534 %
	27-09-2013 al 03-05-2014	23,135 %		27-09-2013 al 03-05-2014	19,270 %
	03-05-2014 al 06-03-2015	18,196 %		03-05-2014 al 06-03-2015	14,838 %

Tabla 5.- Características de la relación en clúster 6 y clúster 7

Las métricas más importantes que se muestran en la tabla 3 y la tabla 5 son: precio de ventas, familia producto y fechas de compra.

La diferencia entre ambas tablas es que en la relación que muestra la tabla 3 se encuentra fuertemente relacionada la métrica familia-producto con el valor de JUGO/NECTAR, mientras que en la tabla 5 se encuentra relacionada la misma métrica, pero con el valor de VINO.

Por lo anterior, podemos deducir que el grupo presentado en la tabla 3 realiza un mayor número de ventas con productos JUGO/NECTAR en ciertos intervalos de tiempo especificados, mientras que el agrupamiento de la tabla 5 muestra un mayor porcentaje de ventas con productos de VINO, en el mismo intervalo de tiempo.

El punto de partida para esta investigación es el análisis detallado que se obtuvo del SQL DATA TOOL con los datos procesados en cada una de las tablas de nuestra base de datos. Las agrupaciones obtenidas en este trabajo fueron comparadas con las agrupaciones obtenidas en la versión paralela, a fin de validar el funcionamiento del algoritmo K-means en la arquitectura CUDA.

Se programó una versión lineal, donde se adaptó la función distancia euclídea en el algoritmo K-means, para contrastar con la versión paralela.

En la tabla 6 se muestra la cantidad de datos procesados con una primera versión paralela del algoritmo K-means, así como los tiempos obtenidos en una versión lineal con la misma cantidad de registros procesados comparado con la versión paralela.

ALGORITMO K-MEANS		
CANTIDAD REGISTROS	VERSION LINEAL TIEMPO	VERSION PARALELA TIEMPO
100	0.010 seg	0.093 seg
1,000	0.253 seg	0.119 seg
10,000	5.790 seg	1.890 seg
1'000,000	17m39.646 seg	3m47.996 seg

Tabla 6.- Resultados obtenidos en tiempo de la versión lineal y versión paralela del algoritmo k-means

Los resultados de la tabla 6 muestran que paralelizar un algoritmo no siempre es lo más viable, ya que procesar pequeñas cantidades de información de forma paralela consume más tiempo en la parte de comunicación entre la CPU y la GPU que el tiempo que se tomaría la CPU en procesarlas.

Sin embargo, la versión paralela muestra una mejora del 80% que la versión lineal en el tiempo de ejecución al procesar 1, 000,000 de registros (como se describe en la tabla 6).

Es importante mencionar que una desventaja significativa al realizar operaciones con números decimales del lado de la GPU es la precisión, ya que una CPU tiene mayor precisión decimal que una GPU.

## V. TRABAJOS FUTUROS

- Migrar la versión paralela del algoritmo K-means al clúster de GPU, utilizando MPI.

- Adaptar al algoritmo K-means otras funciones de distancias, para comparar el desempeño con la función euclídea aplicada al mismo problema.
- Probar con diferentes características, para determinar los agrupamientos en cada centroides determinado para el algoritmo K-means.

## VI. BIBLIOGRAFÍA

- [1] Hernandez Orallo, J., Ramirez, M. J., & Ferri, C. (s.f.). *Introducción a la minería de datos*. PEARSON.
- [2] Pacheco, P. (s.f.). *An Introduction to Parallel Programming*. ELSEVIER -Morgan Kaufmann.
- [3] Cook, S. (s.f.). *CUDA Programming - A developer's guide to parallel computing with GPU's*. ELVESIER - Morgan Kaufmann.
- [4] NVIDIA. (s.f.). *CUDA Parallel Computing Platform*. Obtenido de NVIDIA: [http://www.nvidia.com/object/cuda\\_home\\_new.htm](http://www.nvidia.com/object/cuda_home_new.htm)

# Integración de CUDA C, OpenCV y R, para el procesamiento paralelo de imágenes de satélite

Claudia Martínez Vazquez  
Instituto Tecnológico de La Paz  
La Paz, México  
klauditha84@gmail.com

Saul Martínez Díaz  
Instituto Tecnológico de La Paz  
La Paz, México  
saulmd@itlp.edu.mx

Raul Octavio Martínez Rincon  
CIBNOR S.C.  
La Paz, México  
martinez-rrincon@cibnor.mx

Iliana Castro Liera  
Instituto Tecnológico de La Paz  
La Paz, México  
icastro@itlp.edu.mx

**Resumen**—Debido a los grandes volúmenes de datos que representan las imágenes de satélite una alternativa al software convencional para procesamiento de imágenes de satélite es utilizar una solución a la medida de cómputo paralelo. Este trabajo se enfoca en cómo desarrollar un ambiente de trabajo integrado (procesamiento en paralelo GPGPU - CUDA, librerías especializadas de visión por computadora - OpenCV y pre-procesamiento de las imágenes - R) para el cálculo de un promedio histórico de los índices de vegetación de imágenes de satélite MODIS correspondientes a una región equivalente a la superficie del estado de Baja California Sur, México de un periodo comprendido del año 2000 al 2014, con la finalidad de reducir significativamente los tiempos de procesamiento respecto a los cálculos tradicionales en secuencial desde la CPU.

**Palabras clave**— *CUDA; OpenCV; R; procesamiento de imágenes MODIS.*

## I. INTRODUCCION

El procesamiento de imágenes de satélite es una tarea sumamente demandada y la tendencia es que continúe así, y no es para menos pues sus aplicaciones prácticas involucran tanto el ámbito de la investigación como el empresarial; podemos verlas en detección de incendios en tiempo real, monitoreo de temperatura de los océanos, monitoreo de partículas de aerosoles en la atmósfera, monitoreo de cambios en la vegetación, por mencionar solo algunas.

Actualmente existe una amplia gama de herramientas para análisis espacial ya sea comerciales u open source (ArcGIS, QGIS, MATLAB, R), pero debido a la enorme cantidad de datos y los cálculos intensivos que se realizan, las tareas se ven superadas por elevados tiempos de procesamiento. Es así que una alternativa para abordar este escenario es desarrollar una solución de procesamiento paralelo a la medida, donde el procesamiento se ejecute en un ambiente de trabajo integrado y así la suma de las ventajas individuales de cada tecnología ofrece beneficios significativos en la unión de todas. Este trabajo presenta la integración de CUDA, OpenCV y R con la finalidad de calcular el promedio histórico de los índices de vegetación de diferencia normalizada (por sus siglas en inglés NDVI) de una región equivalente a la superficie del estado de Baja California Sur, México (684 imágenes MODIS de 4800 x 4800 píxeles) del periodo comprendido del año 2000 al 2014.

### A. Trabajos relacionados

Liu, Feld et al. 2015, realizaron dos implementaciones eficientes de un algoritmo de recuperación de Profundidad Óptica de Aerosoles (por sus siglas en inglés AOD), a partir de imágenes de Espectro-radiómetro de Resolución Moderada (por sus siglas en inglés MODIS), empleo dos arquitecturas de cómputo de alto rendimiento diferentes: procesadores de múltiples núcleos y una unidad de procesamiento de gráficos (GPU) [1].

Zhang, You et al. 2011, implementaron técnicas de compresión de datos raster geoespaciales y técnicas de compresión populares. Presentan una aplicación paralela de compresión de datos raster geoespaciales implementada en una GPGPU en comparación con una CPU multi-core con 16 hilos, y consiguen una aceleración de 12x en comparación con la implementación multi-core (con 16 hilos) de la popular librería de compresión Zlib [2].

Zhang, Wang et al. 2010, presentaron soluciones basadas en las ventajas de operaciones en paralelo de una GPGPU; la comparativa consiste en dos algoritmos de procesamiento de imágenes representativos, el detector de bordes de Sobel y filtrado homomórfico. Los datos de las imágenes de prueba son de diferentes resoluciones y se utilizaron tanto en la plataforma de la CPU y la GPU para comparar su eficiencia computacional [3].

Yang, Zhu et al. 2008, hicieron un análisis sobre las distintas características de CUDA GPU y resumen el modo de programa general de CUDA. Además, ponen en práctica varios algoritmos de procesamiento de imágenes clásicos de CUDA, como la ecualización del histograma, la eliminación de las nubes, detección de bordes y codificación/decodificación de transformada discreta del coseno (por sus siglas en inglés DTC) etc., sobre todo introduce los dos primeros algoritmos [4].

## II. CONCEPTOS BASICOS

### A. Imágenes MODIS

Son imágenes generadas con un sensor instalado a bordo del satélite TERRA de la NASA; uno de sus varios productos es denominado MOD13Q1 (en inglés *Vegetation Index 16-Day L3 Global 250m*). El cual contiene los valores de NDVI, los cuales básicamente representan la salud de la cobertura vegetal, es decir qué tan verde o seco se encuentra la superficie del planeta.

Cabe señalar que el formato original de las imágenes MODIS es HDF (por sus siglas en inglés *Hierarchical Data Format*) que es un tipo de empaquetado que contiene a su vez 12 archivos y entre los cuales está el NDVI. Originalmente los valores de los píxeles vienen de -2,000 al 10,000 pero son escalados para manejarse en una escala del -2 al 1, donde valores cercanos a cero representan suelos desnudos, rocas, matorrales o nieve, y cercanos a uno vegetación abundante, como se ilustra en la Fig. 1.

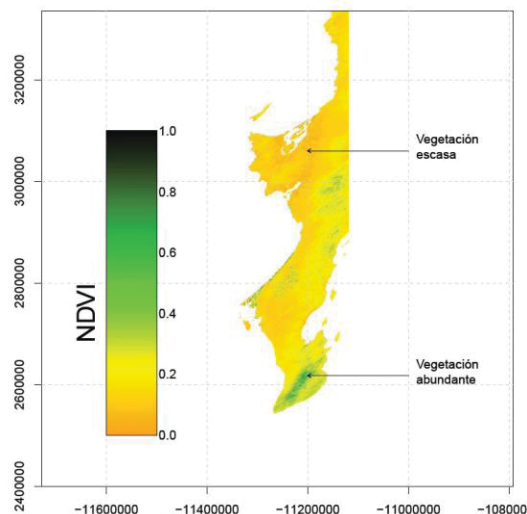


Fig. 1. Valores del NDVI.

Las imágenes MODIS están disponibles desde febrero del año 2000 y en promedio existen de 22 a 23 imágenes por año. Para la serie de tiempo del año 2000 al 2014 existen un total de 342 imágenes, pero para la región correspondiente a Baja California Sur, México, debido a que la superficie completa del estado queda entre dos cuadrantes como se ilustra en la Fig. 2, nos da un total de 684 imágenes a procesar [5].

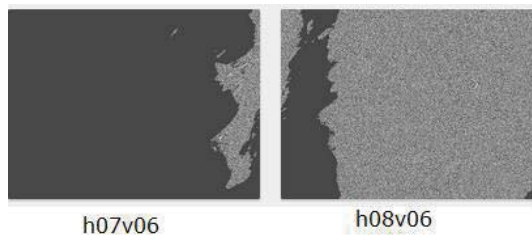


Fig. 2. Cuadrantes MODIS de la ubicación de Baja California Sur.

### B. CUDA

Arquitectura Unificada de Dispositivos de Cómputo (por sus siglas en inglés CUDA), es una plataforma de cómputo paralelo y un modelo de programación inventado por el fabricante de tarjetas de gráficos NVIDIA. Permite aumentos impresionantes en el rendimiento de los cálculos al aprovechar la potencia de la unidad de procesamiento de gráficos (GPU). NVIDIA sabía que un hardware impresionantemente rápido tenía que combinarse con herramientas de hardware y software intuitivas e invitó a Ian Buck a unirse a la compañía y empezar a hacer evolucionar una solución que ejecutara C a la perfección en la GPU. Al reunir el software y el hardware, NVIDIA lanzó CUDA en 2006, la primera solución del mundo para computación general en las GPU [6].

La arquitectura propia de la GPU está constituida por un mayor número de Unidades Aritmético Lógicas (por sus siglas en inglés ALU) respecto a las Unidades de Procesamiento Central (por sus siglas en inglés CPU) esto se ilustra en la Fig.3, por lo que es lógico pensar que los tiempos de procesamiento de tareas que involucren cálculos intensivos se reducen significativamente; esto es posible siempre y cuando las tareas a ejecutar soporten paralelización y la complejidad del cálculo a efectuarse sea importante, debido a que en el modelo de programación de CUDA existe el concepto de *Host* y *Device*, donde el primero se refiere a la CPU y su memoria, y el segundo a la GPU y su memoria. Es así que el envío de información desde el *Host* hacia el *Device* inicia través del *Host*, por lo tanto los tiempos de transferencia de datos entre ambos deben ser menores que el tiempo de ejecución del proceso en la GPU (a través de un *kernel*, los cuales solo viven del lado del *Device* y a su vez son ejecutados por muchos hilos en paralelo en la GPU).

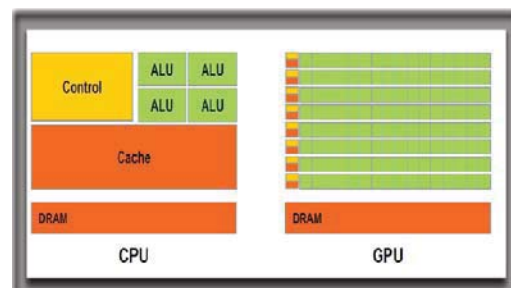


Fig. 3. Esquema de la arquitectura de la CPU y la GPU.

### C. OpenCV

Es una librería de visión por computadora y software de aprendizaje por computadora. OpenCV fue construido para proporcionar una infraestructura común para aplicaciones de visión por computadora y para acelerar el uso de máquinas en los productos comerciales. Al ser un producto de licencia BSD (por sus siglas en inglés *Berkeley Software Distribution*) hace que sea fácil para las empresas utilizar y modificar el código [7].

D. R

R es un lenguaje y entorno de computación y gráficos estadísticos, se compila y se ejecuta en una amplia variedad de plataformas UNIX, Windows y Mac OS; es un proyecto de software libre, que es similar al lenguaje S, que fue desarrollado en los Laboratorios Bell (antes de AT & T, ahora Lucent Technologies) por John Chambers y sus colegas [8].

Entre las ventajas de R podemos contar la gran comunidad activa que lo utiliza, es muy fácil encontrar y reproducir scripts que alguien más hizo sobre un problema y/o tarea en particular, además por ser un lenguaje de programación permite diseñar sus propias funciones y por ser software libre pueden hacer contribuciones o mejorar alguna de las más de 7976 librerías disponibles que ya existen.

III. METODOLOGIA

La solución propuesta se dividió en fases como se ilustra en la Fig.4.

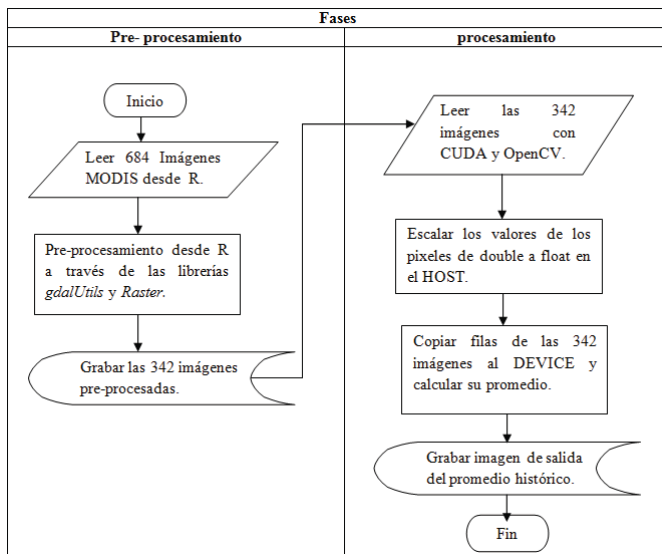


Fig. 4. Diagrama de las fases del procesamiento completo de las imágenes.

A. Pre-procesamiento en R

El procedimiento consiste en la extracción de los datos de NDVI contenido dentro de los archivos HDF y posteriormente recortar la región de interés de cada cuadrante que se ilustran en la Fig.2 quitando las zonas como el Océano Pacífico y el estado de Sonora, posteriormente unirlos y finalizar con el cambio en la proyección, debido a que las imágenes originalmente vienen con proyección Sinusoidal se hace el cambio WGS84.

B. Procesamiento

Una vez generadas las imágenes resultantes del pre-procesamiento en R, se procedió a leerlas en formato TIF para su procesamiento (cálculo del promedio de las 342 imágenes) desde CUDA utilizando la tarjeta NVIDIA que se ilustra en la Tabla I.

Finalmente la lectura se realizó con la librería *highgui.h* (con el objeto *Mat*) de OpenCV, inmerso en código de CUDA C [9]. Una vez leídas las 342 imágenes del lado del *Host*, fue necesario escalar los valores de los píxeles de *double* (8 bytes) a *float* (4 bytes) debido a que no sería suficiente la memoria realmente disponible (8 GB) en el *Device*, y una vez escalados los valores de los píxeles ya de tipo *float* sumaban un total de 7.73 GB de datos a procesar.

TABLA I CARACTERISTICAS TARJETA NVIDIA

Tarjeta de Video		
Modelo	Memoria Global	CUDA Cores
Geforce GTX Titan X	12GB	3072

La estrategia realizada para el envío de los datos fue fraccionar las imágenes y procesarlas por filas, es decir: primero se envió la *fila 1* (1758 columnas de tipo *float*), pero de cada una de las 342 imágenes, es decir se envió una matriz de 342 filas y 1758 columnas, posteriormente se calculó su promedio y se devolvió el resultado hacia el *Host*, y así sucesivamente tantas filas había (3455 filas) en la imagen. La estrategia resulta debido a que las dimensiones originales de las imágenes al ser enviadas al *Device* provocaban violaciones de segmento en la memoria.

El cálculo del promedio en la GPU, se efectuó utilizando 2 bloques con 879 hilos (un total de 1758 hilos) como se bosqueja en la Fig. 5, a cada hilo le correspondió una columna de la matriz de datos recibida (342 filas x 1758 columnas) en el *Device*, el resultado del promedio de la *fila N* recibida se enviaba de regreso al *Host* y se almacenaba en la imagen de salida que finalmente representó el promedio de los 15 años. El envío de datos entre el *Host* y el *Device* se realizó a través de un arreglo bidimensional.

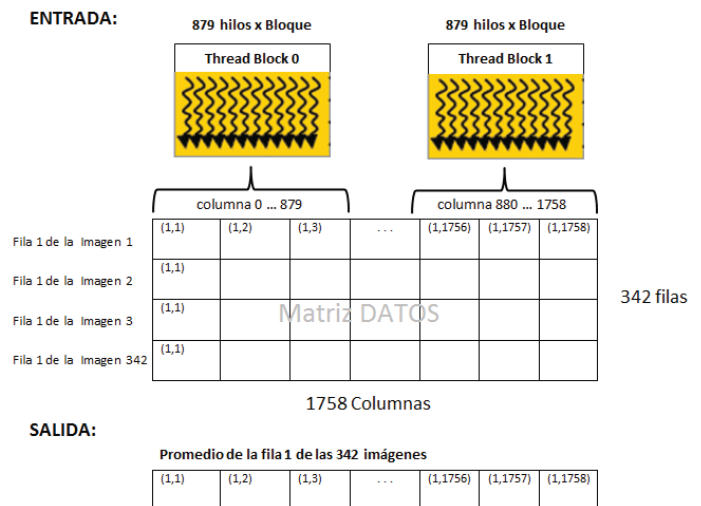


Fig. 5. Gestión de los hilos para procesar la matriz con los datos de la imágenes.

## RESULTADOS

Una vez concluidas las configuraciones que integraron el ambiente de trabajo, al ejecutar el proceso del cálculo del promedio de las 342 imágenes MODIS del NDVI (7.73 GB de datos) implementado en CUDA, resultó ser 67X más rápido que el procedimiento en secuencial desde R en el mismo equipo como se ilustra en la Tabla II. Con esto se confirma la justificación de la tendencia del uso de la GPGPU para cálculos intensivos o procesamiento de grandes volúmenes de datos, por la significativa reducción de tiempos de procesamiento al implementar paralelismo.

TABLA II TIEMPOS PROMEDIO DE EJECUCION

<i>Tiempo (segundos)</i>	
<i>GPU</i>	<i>CPU</i>
75.778	5117.732

## REFERENCIAS

- [1] J. Liu, D. Feld, Y. Xue, S. Member, J. Garcke, and T. Soddemann, "Multicore processors and graphics processing unit accelerators for parallel retrieval of aerosol optical depth from satellite data: implementation, performance, and energy efficiency", IEEE vol.8, pp. 2306–2317, July 2015.
- [2] J. Zhang, S. You, and L. Gruenwald, "Parallel quadtree coding of large-scale raster geospatial data on GPGPUs", Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 457-460, November 2011.
- [3] N. Zhang, J. Wang, and Y. Chen, "Image parallel processing based on GPU". 2nd International Conference on Advanced Computer Control, pp. 367–370, March 2010.
- [4] Z. Yang, Y. Zhu, and Y. Pu, "Parallel Image Processing Based on CUDA". International Conference on Computer Science and Software Engineering, pp. 198–201, December 2008.
- [5] Nvidia (2016), Plataforma de Computación Paralela CUDA, [Online]. Disponible: [http://www.nvidia.com.mx/object/cuda\\_home\\_new\\_la.html](http://www.nvidia.com.mx/object/cuda_home_new_la.html), Accesado: 2016.
- [6] Opencv (2016). ABOUT | OpenCV. [Online]. Disponible: <http://opencv.org/about.html>. Accesado: 2016.
- [7] R-project (2016). R: The R Project for Statistical Computing. [online] Disponible: <https://www.r-project.org/>, Accesado: 2016.
- [8] Guía de instalación R en Linux (2016). The Comprehensive R Archive Network. [Online], Disponible: <https://cran.itam.mx/>, [Accesado: 2016.
- [9] OpenCV en Linux. (2013) Guía de instalación OpenCV en Linux [Online]. Disponible: <https://geekytheory.com/opencv-en-linux/>, Accesado: 2016.

# Diseño de un robot cartesiano para ser empleado en un sistema cartográfico de resolución controlada.

Lic. Ricardo Torres Reyes , Benemérita Universidad Autónoma de Puebla,  
Av. san Claudio y 18 Sur, Ciudad Universitaria, México,  
[ricardoreyes@ece.buap.mx](mailto:ricardoreyes@ece.buap.mx)

Dr. Sergio Vergara Limón, Benemérita Universidad Autónoma de Puebla, Av.  
san Claudio y 18 Sur, Ciudad Universitaria, México, [svergara@ece.buap.mx](mailto:svergara@ece.buap.mx)

Dr. Pedro Javier García Ramírez Investigador T-C. Instituto de Ingeniería  
Boca del Río, Veracruz, México. [jagarcia@uv.mx](mailto:jagarcia@uv.mx)

Dra. María Aurora Diozcora Vargas Treviño, Benemérita Universidad  
Autónoma de Puebla, Av. san Claudio y 18 Sur, Ciudad Universitaria, México,  
[mavargas@ece.buap.mx](mailto:mavargas@ece.buap.mx).

Dra. Amparo palomino merino Benemérita Universidad Autónoma de Puebla,  
Av. san Claudio y 18 Sur, Ciudad Universitaria, México,  
[palomino@ece.buap.mx](mailto:palomino@ece.buap.mx).

**Resumen:** El objetivo de este trabajo es presentar el diseño y modelado de un robot de tipo cartesiano de tres grados de libertad el cual se utilizará para hacer un estudio espacial y puntual de un área de trabajo con una precisión en el orden de micrómetros ( $\mu\text{m}$ ). Se explica el modelo dinámico del robot cartesiano de 3 grados de libertad (gdl). Otros puntos que se aborda en este escrito es una descripción sobre el desarrollo del firmware del sistema, que está dentro de un FPGA de la marca ALTERA de la familia Cyclone II, el cual se encarga de la comunicación de nuestro sistema con la pc, la construcción del robot y los materiales que se utilizaron, también se menciona la caracterización de los motores para que trabajen de forma lineal.

**Abstract—** The main goal of this paper is to present the design and modeling of a Cartesian robot of three degrees of freedom which will be used to make a timely study space and a work area with an accuracy in the order of micrometers ( $\mu\text{m}$ ). The dynamic model is explained. We describe the data acquisition system including a description of the hardware and firmware development. All the system is based on a FPGA of ALTERA Cyclone II family, which is responsible for the communication of our system with the PC. Robot construction and materials used are described and the characterization of a motors to work linearly.

## I. INTRODUCTION

El gran aporte que da la robótica a la medicina es de considerarse, hoy en día es posible realizar cirugías de alto riesgo con una gran seguridad y exactitud, evitando así algún error humano durante los procesos quirúrgicos, realizar cirugías en zonas las cuales el cirujano no puede acceder o visualizar, como son los casos de cirugía de los ojos, a corazón abierto, o en situaciones más complejas como retirar agua del cerebro, además de dañar mucho menos el tejido sano alrededor del área afectada y una recuperación más rápida.

En aplicaciones fisiológicas el poder contar con robots cartesianos de gran resolución permite realizar análisis a gran escala, como es en el caso del campo de los circuitos integrados (CI). Con el continuo crecimiento en tamaño, número de transistores y consumo de potencia eléctrica de los CI, éstos actualmente juegan un papel preponderante en la determinación de la compatibilidad electromagnética de los sistemas electrónicos en los que son utilizados. En general, las estructuras de los CI's así como sus empaquetados tienen poca influencia eléctricamente hablando, y no radian por sí mismos energía de manera eficiente. Por otro lado, las fuentes de Radio Frecuencia (RF) generadas por los CI's, suelen tener que interactuar con una estructura más grande conectada a la placa de un circuito impreso (PCB) para irradiar de manera significativa. Sin embargo, el ruido acoplado derivado de los CI's en estructuras resonantes cerca de un PCB es un problema común en los sistemas digitales de alta velocidad.

## II. DISEÑO DEL ROBOT

Para realizar el diseño del robot cartesiano se utilizó el programa de solidworks ya que es una solución de diseño tridimensional completa que integra un gran número de funciones avanzadas para facilitar el modelado piezas, crear grandes ensamblajes, generar planos y otras funcionalidades que le permiten validar, gestionar y comunicar proyectos de forma rápida, precisa y fiable.

Y para el diseño mecánico se utilizaron Guía de movimiento lineal Tipo HSR como se muestra en la "Figura



1” y las características de esta guía se muestra en la “Tabla 1”.

Tipo estándar



Tipo HSR...A  
(Tipo para carga pesada)

Fig. 1. Guía de movimiento lineal Tipo HSR

Las guías lineales cuadradas son los componentes más profesionales que se pueden usar para realizar movimiento lineal debido a su diseño estructural capaz de soportar grandes cargas y deslizarlas con la menor fricción y juego posible.

Tabla 1: características del guía de movimiento lineal de alta precisión

IGUAL VALOR DE CARGA EN LOS 4 SENTIDOS
ALTA RIGIDEZ
CAPACIDAD DE AUTO-AJUSTE
EXCELENTE DURABILIDAD

Se utilizaron carritos de alta precisión para poder producir los movimientos lineales “Figura 2”.



Fig. 2. Carro de precisión.

En la “Tabla 2” se muestran las características del carro de precisión

MÁS BAJO NIVEL DE RUIDO DE FUNCIONAMIENTO
RETENCIÓN DE GRASA ÓPTIMA DENTRO DEL BLOQUE DE GUÍA.

VIDA ÚTIL MÁS LARGA.
BOLA ESPACIADOR DISEÑADO PARA DURAR TODA LA VIDA ÚTIL DEL COJINETE.

Se utilizaron tornillos embalados el cual se muestra en la “Figura 3” ya que es un eje roscado que proporciona un camino de rodadura helicoidal a unos rodamientos de bolas que actúan como un tornillo de precisión. Como el movimiento se realiza por rotación, no por deslizamiento, el rozamiento es menor, la fricción del conjunto es baja. Y como el esfuerzo se reparte entre varias bolas, es capaz de aplicar o resistir altas cargas de empuje. Se pueden fabricar con unas tolerancias estrechas y por tanto son adecuados para su empleo donde se necesita una alta precisión.



Fig. 3. tornillos embalados

Diseño y construcción del eje Z

En la “Figura 4”, se puede apreciar la estructura mecánica del robot y las partes que la componen para su diseño, la configuración del robot es de tipo cartesiano, por lo que, se debe desplazar linealmente en el eje Z.

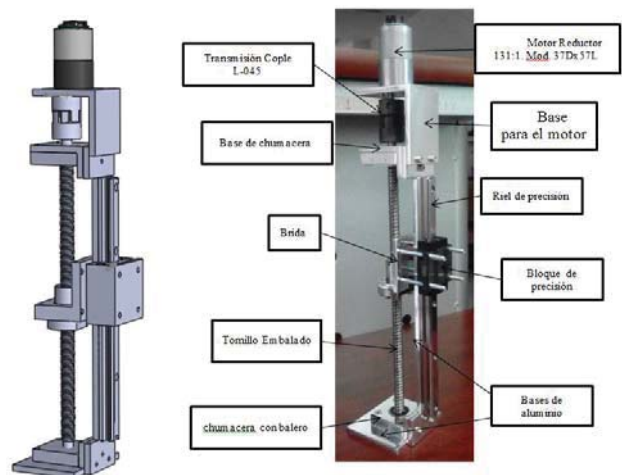


Fig. 4. Estructura mecánica del eje Z

Diseño y construcción mecánica del eje X

El eje X cuenta con un desplazamiento de 300 milímetros (mm) ver “Figura 5” se puede apreciar la estructura mecánica del robot y las partes que la componen como también se puede ver su diseño en Solidworks la configuración del robot es de tipo cartesiano, por lo que, se debe desplazar linealmente en el eje X.

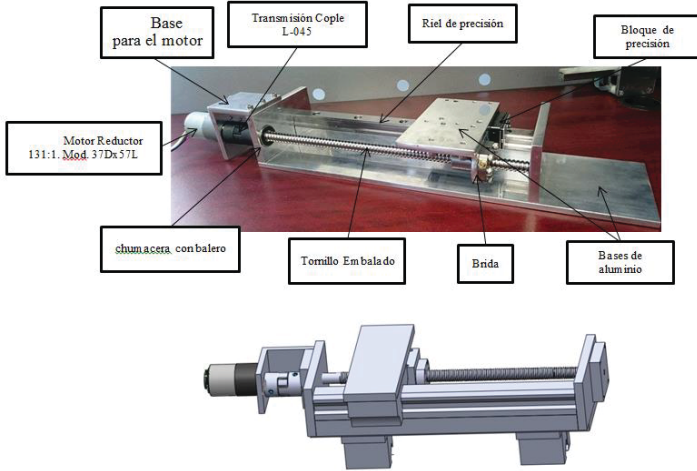


Fig. 5. Estructura mecánica del eje x

Diseño y construcción mecánica del eje Y

Para el último eje se utilizó una cremallera y un piñón ya que el mecanismo de piñón y cremallera permite transformar el movimiento circular en rectilíneo como también puede transformar el movimiento rectilíneo en movimiento circular. El eje Y, cuenta con un desplazamiento de 100 milímetros (mm) ver “Figura 6” se puede apreciar la estructura mecánica del eje y las partes que la componen como también se puede ver su diseño en Solidworks , la configuración del robot es de tipo cartesiano, por lo que, se debe desplazar linealmente en el eje Y.

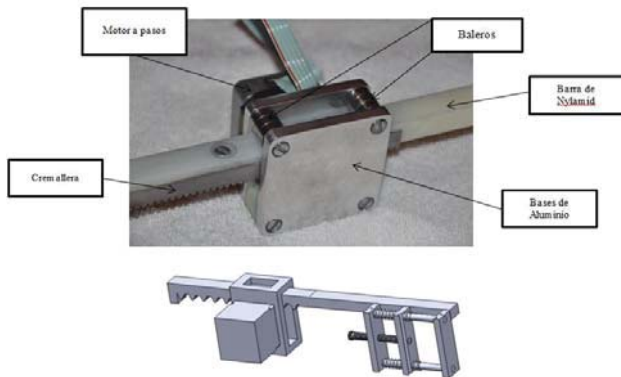


Fig. 6. Estructura mecánica del eje x

III. CINEMÁTICA Y DINÁMICA

Modelo dinámico del sistema.

Debemos de conocer el modelo para estudiar al robot de manera correcta, y en esta deben estar presentes los fenómenos físicos como son los pares gravitacionales, las fuerzas centrípetas y de coriolis, así como la fricción de nuestro sistema. Con el modelo dinámico se pretende crear una regla matemática que relacione las variables de entrada y salida de nuestro sistema.

La cinemática se encarga de estudiar el movimiento del robot con respecto al sistema de referencia, permitiendo la identificación de la posición y orientación del elemento final del robot sin considerar las fuerzas o momentos que originan el movimiento. La cinemática directa se encarga de determinar la posición del elemento final con respecto al sistema de referencia [4].

En otras palabras, el modelo cinemático directo, describe la relación entre la posición articular y la posición y orientación x del dispositivo terminal del robot, es decir, el modelo cinemático directo de un robot es una relación de la forma.

$$x = f(q) \tag{1}$$

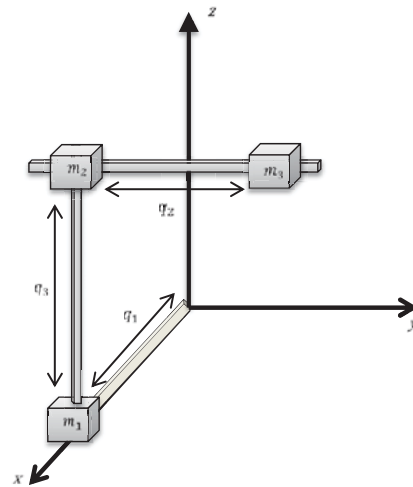


Fig. 7. Diagrama robot cartesiano de 3 g.d.l.

El modelo cinemático directo del robot cartesiano de acuerdo a la figura anterior queda representado de la siguiente forma:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} q_1 \\ 0 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ q_2 \\ 0 \end{bmatrix}; \quad \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ q_2 \end{bmatrix} \tag{2}$$

Modelo Dinámico Uno de los procedimientos más empleados para la obtención de modelos dinámicos de los robots, en forma cerrada, es el basado en las ecuaciones de movimiento de Euler Lagrange. El empleo de las ecuaciones de Euler Lagrange para el modelado, requiere de dos conceptos importantes: La energía cinética y la energía potencial. [7]

La energía total de un robot manipulador de n.g.l. es la suma de sus energías cinética y potencial:

$$\varepsilon(q, \dot{q}) = K(q, \dot{q}) + U(q) \tag{3}$$

El lagrangiano de un robot de n.g.l. es la diferencia entre su energía cinética y su energía potencial:

$$L(q, \dot{q}) = K(q, \dot{q}) - U(q) \quad (4)$$

Las ecuaciones de movimiento de Euler- Lagrange para un manipulador de n.g.l, vienen dadas por:

$$\tau = \frac{d}{dt} \left[ \frac{\partial L(q, \dot{q})}{\partial \dot{q}} \right] - \left[ \frac{\partial L(q, \dot{q})}{\partial q} \right] \quad (5)$$

El modelo cinemático directo del robot cartesiano de acuerdo a la "Figura 7" queda representado de la siguiente forma:

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} q_1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ q_2 \\ 0 \end{bmatrix} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ q_3 \end{bmatrix} \quad (6)$$

La energía cinética total del robot viene dada por:

$$K(q, \dot{q}) = \frac{1}{2} m v^2 \quad (7)$$

Donde la velocidad es la derivada de la posición

$$v = \frac{d}{dt} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (8)$$

$$v_1 = \frac{d}{dt} \begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = \begin{bmatrix} \dot{q}_1 \\ 0 \\ 0 \end{bmatrix} \quad (9)$$

$$v_2 = \frac{d}{dt} \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = \begin{bmatrix} 0 \\ \dot{q}_2 \\ 0 \end{bmatrix} \quad (10)$$

$$v_3 = \frac{d}{dt} \begin{bmatrix} x_3 \\ y_3 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \dot{q}_3 \end{bmatrix} \quad (11)$$

Elevando al cuadrado nos queda de la siguiente manera.

$$v^2 = \|v\|^2 = v^T v = [V_1 \quad V_2 \quad V_3] \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \quad (12)$$

$$v_1^2 = \|v_1\|^2 = [\dot{q}_1 \quad 0 \quad 0] \begin{bmatrix} \dot{q}_1 \\ 0 \\ 0 \end{bmatrix} = \dot{q}_1^2 \quad (13)$$

$$v_2^2 = \|v_2\|^2 = [0 \quad \dot{q}_2 \quad 0] \begin{bmatrix} 0 \\ \dot{q}_2 \\ 0 \end{bmatrix} = \dot{q}_2^2 \quad (14)$$

$$v_3^2 = \|v_3\|^2 = [0 \quad 0 \quad \dot{q}_3] \begin{bmatrix} 0 \\ 0 \\ \dot{q}_3 \end{bmatrix} = \dot{q}_3^2 \quad (15)$$

Ahora que tenemos las ecuaciones de las velocidades obtenemos las ecuaciones de la energía cinética para cada grado de libertad.

$$K_1(q, \dot{q}) = \frac{m_1 \dot{q}_1^2}{2}, \quad (16)$$

$$K_2(q, \dot{q}) = \frac{m_2 (\dot{q}_1^2 + \dot{q}_2^2)}{2}, \quad (17)$$

$$K_3(q, \dot{q}) = \frac{m_3 (\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2)}{2}, \quad (18)$$

Sumando todas estas energías tenemos que la energía cinética total está dada como:

$$K_1(q, \dot{q}) = \frac{m_1 \dot{q}_1^2}{2} + \frac{m_2 (\dot{q}_1^2 + \dot{q}_2^2)}{2} + \frac{m_3 (\dot{q}_1^2 + \dot{q}_2^2 + \dot{q}_3^2)}{2} \quad (19)$$

Aplicando algebra queda de la siguiente manera:

$$K(q, \dot{q}) = \frac{1}{2} m_1 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_1^2 + \frac{1}{2} m_2 \dot{q}_2^2 + \frac{1}{2} m_3 \dot{q}_1^2 + \frac{1}{2} m_3 \dot{q}_2^2 + \frac{1}{2} m_3 \dot{q}_3^2 \quad (20)$$

Factorizamos la ecuación:

$$K(q, \dot{q}) = \frac{1}{2} [m_1 + m_2 + m_3] \dot{q}_1^2 + \frac{1}{2} [m_2 + m_3] \dot{q}_2^2 + \frac{1}{2} [m_3] \dot{q}_3^2 \quad (21)$$

Energía potencial

$$U(q) = m_3 g q_3 \quad (22)$$

De las ecuaciones encontradas obtenemos el lagrangiano:

$$L(q, \dot{q}) = K(q, \dot{q}) - U(q) \quad (23)$$

$$L(q, \dot{q}) = \frac{1}{2} [m_1 + m_2 + m_3] \dot{q}_1^2 + \frac{1}{2} [m_2 + m_3] \dot{q}_2^2 + \frac{1}{2} [m_3] \dot{q}_3^2 - m_3 g q_3 \quad (24)$$

Ahora encontramos las ecuaciones de movimiento con la siguiente ecuación

$$\frac{d}{dt} \left[ \frac{\partial L(q, \dot{q})}{\partial \dot{q}} \right] - \left[ \frac{\partial L(q, \dot{q})}{\partial q} \right] \quad (25)$$

$$\begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \end{bmatrix} = \begin{bmatrix} m_1 + m_2 + m_3 & 0 & 0 \\ 0 & m_2 + m_3 & 0 \\ 0 & 0 & m_3 \end{bmatrix} \begin{bmatrix} \ddot{q}_1 \\ \ddot{q}_2 \\ \ddot{q}_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ m_3 \end{bmatrix} g \quad (26)$$

$$\tau_1 = (m_1 + m_2 + m_3) \ddot{q}_1 \quad (27)$$

$$\tau_2 = (m_2 + m_3) \ddot{q}_2 \quad (28)$$

$$\tau_3 = (m_3) \ddot{q}_3 + (m_3) g \quad (29)$$

Donde  $\tau_1$ ,  $\tau_2$  y  $\tau_3$  son las fuerzas aplicadas.

#### IV. DISEÑO DEL HARDWARE.

En nuestra electrónica se ocupó un FPGA, para realizar la adquisición de datos de nuestro sistema, esta tarjeta fue desarrollada en la facultad de ciencias de la electrónica (FCE) en la Benemérita Universidad Autónoma de Puebla (BUAP).

Una parte fundamental del control del sistema, son los actuadores en este caso nuestro sistema contara con dos actuadores para poder desplazar los ejes X y Z de nuestro robot cartesiano en este caso la elección de nuestros actuadores son motores 37DX54L “Figura 8”, ya que estos motores nos proporcionan un buen par para poder ofrecer una buena movilidad de nuestro sistema. Estos motores cuentan con encoders del tipo incremental los cuales servirán como sensores para retroalimentar nuestro sistema de control del sistema, este encoder está basado sensores Hall, detecta la rotación del eje del motor.

El sensor Hall requiere un voltaje ( $V_{cc}$ ) de entre 3.5V y 20V y una corriente máxima de 10mA. Las salidas A y B son ondas cuadradas desde 0V hasta  $V_{cc}$  con un desfase de 90°C. La frecuencia de las transiciones te dice la velocidad del motor y el orden te da la dirección. [8]

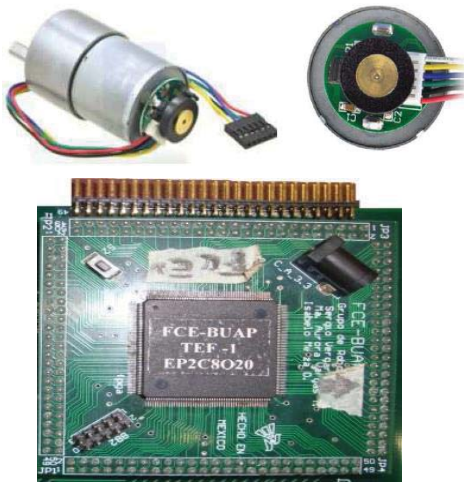


Fig. 8. Sistema de hardware

Para controlar el guiro de nuestros motores se utiliza un puente H, IBT-2

El cual recibe la señal enviada desde el FPGA para realizar la activación de los motores esta tarjeta también controlara el giro recibiendo una señal desde el FPGA y la mandara a nuestros actuadores, en este caso nuestros dos motores esto con el fin de tener un buen acondicionamiento de nuestras señales enviadas del FPGA hasta nuestros sistemas electrónicos de potencia (Figura 9).

El Puente H que se utilizara para controlar el eje X y el eje Z es el IBT-2 de Arduino.

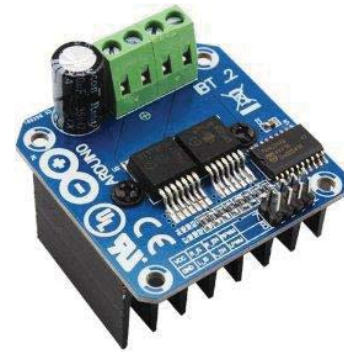


Fig. 9. Puente H IBT-2 de Arduino.

#### Características

Soporta una tensión de alimentación de 5.5V a 27V.

Una Corriente máxima de 43.

Nivel de entrada de 3.3V -5V.

Contiene 8 puertos de entrada

1. RPWM se activa en alto gira hacia adelante
2. LPWM se activa en alto gira al reverso
3. R\_EN enable se activa en alto
4. L\_EN enable se activa en alto
5. VCC voltaje de entrada 3.3V -5V
6. GND es la tierra

#### Diseño del Firmware.

Para la programación del firmware se ocupará el software Quartus II versión 13.0 de altera y un FPGA EP2C8Q208C7 de la familia Cyclone II. Todo esto se realizará mediante el código de programación AHDL, el cual controlara todos nuestros circuitos electrónicos.

Para poder realizar la lectura de nuestros encoders, se desarrollará un bloque para realizar dicha lectura, como ya se había mencionado con anterioridad los encoders utilizados en nuestro sistema son del tipo incremental, en base a una metodología para poder saber la posición del motor de manera incremental se realizó un código para visualizar las lecturas enviadas por los encoder lo cual nos generó un bloque de firmware. [9]

Con este bloque se obtiene el conteo de los encoders incrementales con toda precisión, sin perder cuentas y garantizando el máximo desempeño de nuestro sistema.

También se ocupará una metodología de modulación por ancho de pulsos ya que es una técnica utilizada para regular la velocidad de giro de los motores eléctricos de inducción o asíncronos. Mantiene el par motor constante y no supone un desaprovechamiento de la energía eléctrica. Se utiliza tanto en corriente continua como en alterna, como su nombre lo indica, al controlar: un momento alto (encendido o alimentado) y un momento bajo (apagado o desconectado). [9]

Partiendo de la metodología de la modulación por ancho de pulsos, se crea el código para poder generar ese tipo de modulación, ya que para poder realizar el control de par de los actuadores los generaremos a través de esta técnica de modulación por anchos de pulsos, conociendo como funciona esta metodología se desarrolla un bloque de generación de ancho de pulsos.

V. DISEÑO DEL SOFTWARE.

En los apartados anteriores se describe las partes que realizan el control de nuestro robot cartesiano, como son el hardware y el firmware que nos brindara un buen control para nuestro sistema.

En el siguiente diagrama a bloques “Figura 10” se puede observar la conexión de las etapas de nuestro sistema, que se requiere para nuestro experimento, podemos observar en donde van conectados nuestros motores, encoders y los canales para las señales recibidas del experimento, cada motor cuenta con dos señales que hay que manipular para nuestro sistema señal (A y B), las cuales deben de ser acopladas de manera correcta a nuestro FPGA, ya que cualquier pico de voltaje podría dañar nuestro FPGA, diseñada previamente en la facultad de ciencias de la Electrónica (BUAP), esta tarjeta es para poder enviar y recibir datos de la computadora al FPGA y así poder hacer el proceso de nuestras señales de los encoder de los motores en nuestra computadora, una vez procesadas la señales con el ordenador, con la ayuda del software LABVIEW, se controlaran todos esas señales y se podrán visualizar en la pantalla del usuario para que este pueda darle órdenes a nuestro robot cartesiano y así poder realizar el control del sistema, de tal manera que pueda ser entendible para el usuario que quiera realizar un tarea en el sistema.

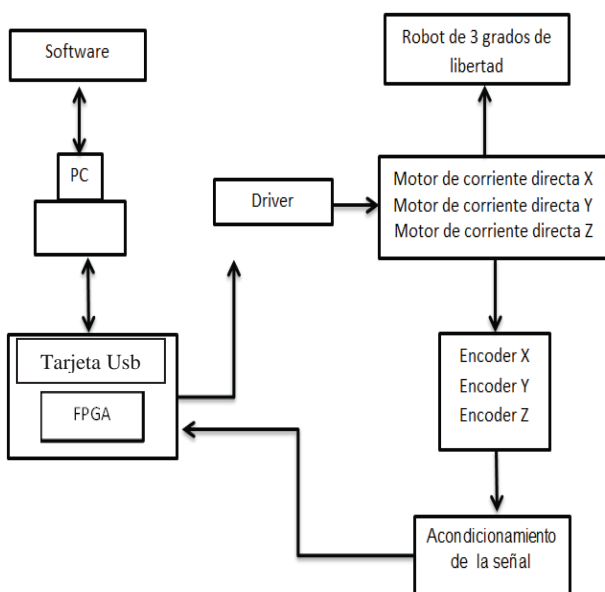


Fig. 10. Diagrama a bloques del sistema

voltaje	Corriente	Kilogramos(kg)	Potencia	Torque	Eficiencia
3	0.2	0.43	0.6	0.16071723	0.26786205
4	0.64	1.52	2.56	0.56811672	0.22192059
5	1.03	2.2	5.15	0.8222742	0.15966489
6	1.39	2.83	8.34	1.05774363	0.12682777
7	1.7	3.7	11.9	1.3829157	0.1162114
8	2.05	4.28	16.4	1.59969708	0.0975425
9	2.45	4.65	22.05	1.73798865	0.07882035
10	2.63	5	26.3	1.868805	0.07105722
11	2.97	5.6	32.67	2.0930616	0.06406678
12	2.99	6	35.88	2.242566	0.06250184
13	2.85	6	37.05	2.242566	0.0605281

Una de las partes importantes en nuestro sistema es la parte del control ya que una vez obtenidos los parámetros del encoder, el error del sistema, las posiciones, las velocidades, se puede ingresar todo a la ecuación de control, en este caso utilizaremos la tangente hiperbólica como controlador, ya que tiene un buen desempeño [1], todo esto utilizando una estructura básica de sistema de control como se puede observar en la “Figura 11”.

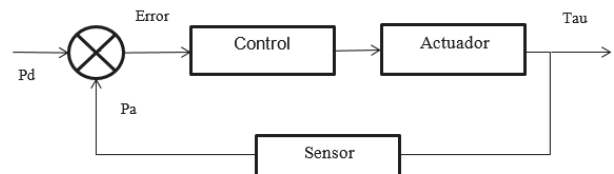


Fig. 11. Diagrama de control en lazo cerrado

$$\tau = K_p * \tanh(E) - K_v * \tanh(\text{vel})$$

Dónde:

**Kp y Kv** = Son las ganancias para el desempeño del controlador.

**E** = Error del sistema.

**vel** = Velocidad con la que se desplaza el eje en el riel  $\frac{mm}{s^2}$ .

**τ** = Es la energía proporcionada por el controlador en este caso es torque **Nm**.

VI. RESULTADOS EXPERIMENTALES.

Caracterización del primer motor de CD

Para caracterizar el motor se necesitamos saber el voltaje nominal al que trabaja nuestro motor esto se obtiene variando

el voltaje desde 3v a 12v, midiendo la corriente y los kilos que el motor llega a levantar posteriormente teniendo esos datos sacamos la potencia, el torque y la eficiencia como se muestra en la “Tabla 3”.

Tabla 3. Datos del Voltaje Nominal

Después de haber realizado las mediciones correspondientes realizamos una gráfica de eficiencia contra voltaje en el cual se puede observar que nuestro motor tiene un voltaje nominal de 12v ya que tiene una eficiencia de 0.06250184 y un torque de 2.242566.

Posteriormente se encuentra la frecuencia de 453 Hz, ya que es la menos audible y es con la que trabaja mejor el motor, ya que para las otras frecuencias al variar el ciclo de trabajo no trabajaban como se esperaba.

Una vez elegida la frecuencia a trabajar, se varía el ancho de pulso o ciclo de trabajo del 10% hasta el 90%, esta variación nos entregara distintas fuerzas, las cuales nos serán de mucha utilidad para crear nuestro control como se muestra en la “Tabla 4”.

Tabla 4 valores del ciclo de trabajo

Voltaje nominal: 12v		
Frecuencia: 453Hz	kilos(kg)	Torque
10%	0	0
20%	0.575	0.21491258
30%	1.475	0.55129748
40%	2.175	0.81293018
50%	3.05	1.13997105
60%	3.56	1.33058916
70%	4.115	1.53802652
80%	5.035	1.88188664
90%	5.65	2.11174965

Una vez que se encontraron las fuerzas con esta frecuencia la graficamos para ver si la respuesta del motor es lineal, lo que el control pida sera lo que se le entregue. “Figura 12”.

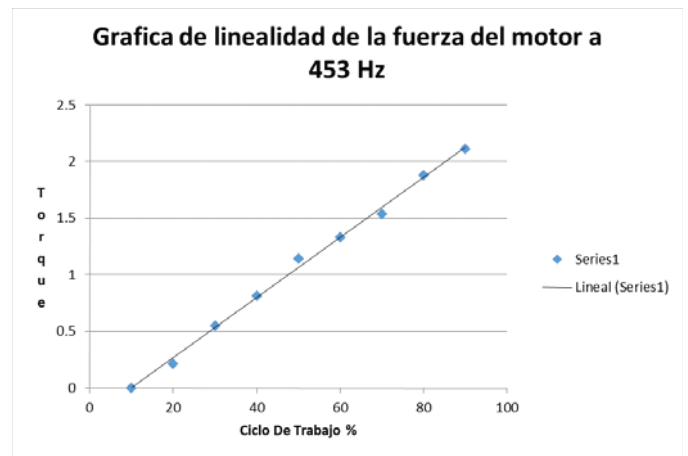


Fig. 12. Linealidad de la fuerza del motor

Caracterización del segundo motor

Para el segundo motor se utiliza el mismo procedimiento que es saber el voltaje nominal al que trabaja nuestro motor esto se obtiene variando el voltaje desde 3v a 12v, midiendo la corriente y los kilos que el motor llega a levantar posteriormente teniendo esos datos sacamos la potencia, el torque y la eficiencia como se muestra en la “Tabla 5”.

Tabla 5. Datos del Voltaje Nominal

voltaje	Corriente	kilogramos(kg)	Potencia	Torque	Eficiencia
3	0.41	0.51	1.23	0.19061811	0.15497407
4	0.7	1.15	2.8	0.42982515	0.15350898
5	1.02	2	5.1	0.747522	0.14657294
6	1.32	2.61	7.92	0.97551621	0.12317124
7	1.64	3.41	11.48	1.27452501	0.11102134
8	1.96	4.11	15.68	1.53615771	0.09796924
9	2.21	4.75	19.89	1.77536475	0.08925916
10	2.45	5.44	24.5	2.03325984	0.0829902
11	2.75	5.95	30.25	2.22387795	0.07351663
12	2.9	6.5	34.8	2.4294465	0.06981168
13	3.1	7.08	40.3	2.64622788	0.06566322

Después de haber realizado las mediciones correspondientes realizamos una gráfica de eficiencia contra voltaje en el cual se puede observar que nuestro motor tiene un voltaje nominal de 12v ya que tiene una eficiencia de 0.06981168 y un torque de 2.4294465.

posteriormente se encuentra la frecuencia de 200 Hz, ya que es la menos audible y es con la que trabaja mejor el motor, ya que para las otras frecuencias al variar el ciclo de trabajo no trabajaban como se esperaba.

Una vez elegida la frecuencia a trabajar, se varía el ancho de pulso o ciclo de trabajo del 10% hasta el 90%, esta variación nos entregara distintas fuerzas, las cuales nos serán de mucha utilidad para crear nuestro control como se muestra en la “Tabla 6”.

Tabla 6 valores del ciclo de trabajo

Voltaje nominal: 12v		
Frecuencia: 200Hz	kilos(kg)	Torque
10	0.035	0.01308164
20	0.84	0.31395924
30	1.35	0.50457735
40	2.16	0.80732376
50	3.5	1.3081635
60	4.14	1.54737054
70	4.9	1.8314289
80	5.1	1.9061811
90	5.58	2.08558638

Una vez que se encontraron las fuerzas con esta frecuencia la graficamos para ver si la respuesta del motor es lineal, lo que el control pida sera lo que se le entregue. “Figura 13”.

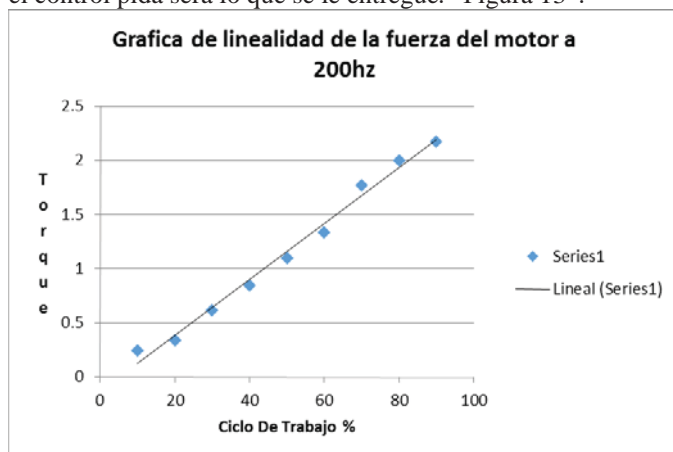


Fig. 13. Linealidad de la fuerza del motor

El robot de tres grados de libertad en configuración cartesiana se desarrolla por la necesidad de tener un dispositivo capaz de hacer un estudio espacial puntual con gran precisión. Se plantea de propósito general ya que no sólo se pretende usar

en una sola aplicación, si no que puede ser modular. Una de las otras aplicaciones que se le desea dar a este robot, es para realizar el sensado de variables fisiológicas en mamíferos pequeños, por lo cual se necesita que este dispositivo sea capaz de hacer movimientos en las coordenadas “x”, ”y” y “z” con gran precisión y exactitud.

Las especificaciones técnicas se muestran en la “Tabla 7” y en “Figura 14” se muestra el sistema terminado

Tabla 7 Especificaciones Técnicas

Estructura mecánica:	Tipo Cartesiano
Grados de libertad (GDL):	3 GDL
Rango de movimiento:	Eje Z (vertical) , eje X y Y(horizontal)
Realimentación:	Encoder de cuadratura con una resolución de 64 pulsos por revolución del eje del motor.
Actuadores:	-Motor de C.D. 12VCD con una caja de engranajes de metal 131.25:1.
Capacidad de los motores:	80 RPM y 300 mA funcionando libremente, 250 oz-in (18 kg-cm) y hasta 5A puesto.
Transmisión:	Copleé L-045
Precisión en el movimiento lineal para el eje Z :	-Riel para guía lineal de 15mm -Bloque lineal para Guía de 15mm. -Tornillo embalado C/Tuerca de 12 mm.
Precisión en el movimiento lineal para el eje X :	-Riel para guía lineal de 20mm -Bloque lineal para Guía de 20mm. -Tornillo embalado C/Tuerca de 12 mm.
Precisión en el movimiento lineal para el eje Y :	-Cremallera -Piñón. Motor a pasos
Peso del eje X	3.555kg
Peso del eje Z	2.005kg
Peso del eje Y	705 gr
Peso total del sistema	6.265kg

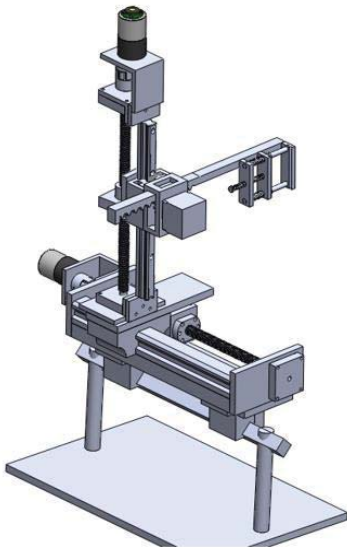


Fig. 14. Prototipo del robot cartesiano.

## VII. CONCLUSIONES Y TRABAJO A FUTURO

En nuestro sistema mecánico se puede observar que, con ayuda de las guías de movimiento lineal, carros de precisión y tornillos embalados podremos alcanzar la exactitud que queremos conseguir, una precisión en el orden de las micras gracias a la ayuda de los motores con reductores podremos alcanzar la precisión antes mencionada. Se obtuvo el modelo cinemático y dinámico de nuestro robot de 3 gdl, en base a este se puede predecir con exactitud el comportamiento del sistema realizando simulaciones y observar los pares necesarios para el movimiento del sistema. También se caracterizaron los motores de los ejes X y Z ya que este motor tiene que trabajar de manera lineal porque cuando el control le pida el torque necesario, el motor dará el torque que se necesita para poder mover nuestro eje.

Como trabajo a futuro se acabará de desarrollar el firmware para que se acoplen todas las etapas requeridas para su buen funcionamiento y posteriormente se realicen las pruebas experimentales en el área de investigación biológica encontramos sistemas estereotáxicos (un tipo de robot cartesiano) para animales tales como ratas o ratones. Este tipo de dispositivos estereotáxicos permiten hacer estudio de campos magnéticos biológicos los cuales tienen su origen en corrientes eléctricas que circulan en algunas células, como en el sistema nervioso y en el corazón, o en materiales magnéticos acumulados en ciertos órganos, como en el hígado y los pulmones. Medir tales campos permite localizar la región que los produce y determinar la intensidad de la corriente o la concentración de los materiales magnéticos acumulados.

## Referencias.

- [1] Vergara S., Vargas M.A., Paic G., Tejeda G., Fernandez A., León I., Reyes F., and Villasenor L. Characterization of the ACORDE scintillator counters using a PCI electronic card, volume 53 (2) 120-125. Revista Mexicana de Física, México, Abril 2007.
- [2] Dorian Rojas Balbuena, Diseño, construcción e instrumentación de un robot cartesiano pick and place de 4GDL, Tesis de maestría, Puebla, Puebla, junio del 2011.
- [3] Josep Balcells, José Luis Romeral, Autómatas Programables, marcombo, 1997.
- [4] Fernando Reyes Cortes, Robótica control de robots manipuladores, alfa omega, 1er edición
- [5] Shanley T. and Don A. PCI System Architecture. AddisonWesley Developer's Press, fourth edition, 1999
- [6] Javier Vargas Valencia, Iliana Ramírez Velázquez, Santiago Pérez Walton, física mecánica conceptos básicos y problemas, Editorial ITM, 1a edición diciembre del 2008
- [7] Control de Movimiento de Robots Manipuladores, Rafael Kelly, Editorial PEARSON Prentice Hall.
- [8] Josep Balcells, José Luis Romeral, Autómatas Programables, marcombo, 1997.
- [9] Frank A. Scarpino, Ph.D., Frank Scarpino, Vhdl and Ahdl Digital System Implementation, illustrated, Prentice Hall International, 1998.



# Diseño e instrumentación de un sistema de recepción OFDMA en plataforma FPGA.

J.C. Gutiérrez Ortega, J. Castañeda Camacho, S. Vergara Limon, J. E. M. Gutiérrez Arias.  
 Facultad de Ciencias de la Electrónica (FCE), Maestría en Ciencias de la Electrónica Opción en Automatización,  
 Benemérita Universidad Autónoma de Puebla (BUAP), Av. San Claudio y 18 Sur S/N. C.U.,  
 Edificio 109A C.P. 72570 Puebla, Pue., México.

**Resumen:** El Multiplexaje por División Ortogonal de Frecuencia (OFDM) es una técnica de modulación popular y ampliamente aceptada en el área de comunicaciones inalámbricas. La versión multiusuario del sistema OFDM es llamada Acceso Múltiple por División Ortogonal de Frecuencia (OFDMA). La principal ventaja de OFDMA es su robustez al desvanecimiento del canal en el entorno inalámbrico. En un sistema OFDMA la Transformada Rápida de Fourier (FFT) y su contraparte la transformada rápida de Fourier inversa (IFFT) se utilizan para modular y demodular los símbolos de datos sobre las subportadoras. En OFDMA la idea es utilizar un número de subportadoras esparcidas uniformemente sobre una banda de frecuencia, de tal manera que el ancho de banda disponible se utilice para una máxima eficiencia. OFDM y OFDMA tienen muchas ventajas que contribuyen a su popularidad; por lo tanto, ha sido usado en muchos sistemas inalámbricos y adoptado por diversas normas. El estándar IEEE 802.16e se basa en el empleo de OFDMA y ha surgido como un fuerte candidato para futuros sistemas inalámbricos. El objetivo de este trabajo es implementar un sistema OFDMA basado en el estándar IEEE 802.16e usando un FPGA de Altera de la familia Cyclone III y mostrar los resultados de las simulaciones de los bloques utilizados en el sistema propuesto usando Matlab y Quartus II.

**Abstract:** Orthogonal Frequency Division Multiplexing (OFDM) is a popular and widely accepted modulation technique in the area of wireless communication. Multi-user version of the OFDM system is called Orthogonal Frequency Division Multiple Access (OFDMA). The main advantage of OFDMA is its robustness to channel fading in wireless environment. In the OFDMA systems the Fast Fourier Transform (FFT) and its counterpart the Fast Fourier Transform Inverse (IFFT) are used to modulate and demodulate the data symbol on the subcarriers. In OFDMA the idea is to utilize a number of subcarriers, spread regularly over a frequency band, in such a way so that the available bandwidth is utilized to maximal efficiency. OFDM and OFDMA have many benefits contributing to its popularity; therefore, it has been used in many wireless systems and adopted by various standards. The standard IEEE 802.16e uses an OFDMA access technique and it has emerged as a strong candidate for future wireless systems. The objective of this paper is the implementation of the OFDMA system based on the standard IEEE 802.16e using a FPGA Altera Cyclone III family and show results of simulations of the blocks used in the proposed system by using Matlab and QuartusII.

**Keywords:** FFT, IFFT, FPGA, OFDMA, OFDM.

## I. INTRODUCCIÓN

Hoy en día la cantidad de usuarios de los sistemas inalámbricos ha ido en aumento y no solo eso, el incremento

del requerimiento de mayores tasas de datos ha crecido drásticamente haciendo de gran importancia la renovación de los sistemas de comunicaciones para poder satisfacer dichas demandas. Una de las propuestas tecnológicas que ha tenido gran aceptación está basada en el uso del Multiplexaje por División Ortogonal de Frecuencia que permite la transmisión de grandes cantidades de datos digitales, a través de una onda de radio. Este esquema ha sido usado en Europa para la transmisión de televisión y radio digital, pero a nivel mundial la atención hacia OFDM y su versión multiusuario OFDMA ha ido en aumento haciendo de estos claramente la tendencia a seguir para los próximos años [1].

De manera conceptual OFDM ha existido durante décadas, pero su implementación real y con costos aceptables ha sido posible desde la llegada y propagación de tecnologías como los microprocesadores de alta velocidad y los dispositivos lógico programables para poder hacer fiable el procesamiento digital requerido [2].

La implementación de este esquema de acceso requiere de la IFFT para lograr la transmisión por medio de subportadoras ortogonales, por eso el número de subportadoras está ligado al número de muestras de la transformada. De manera general, OFDM se refiere a la transmisión de una trama digital que requiere una gran tasa de transferencia mediante  $N$  líneas paralelas más lentas en subportadoras contiguas y ortogonales que transportan símbolos independientes que son producto de algún tipo de modulación digital como QPSK, 16-QAM, 64-QAM [3]. La Figura 1 muestra un ejemplo con el módulo de los espectros correspondientes a un conjunto de 6 subportadoras. La diferencia entre OFDM y OFDMA es que en la última cada subportadora puede pertenecer a usuarios distintos [4].

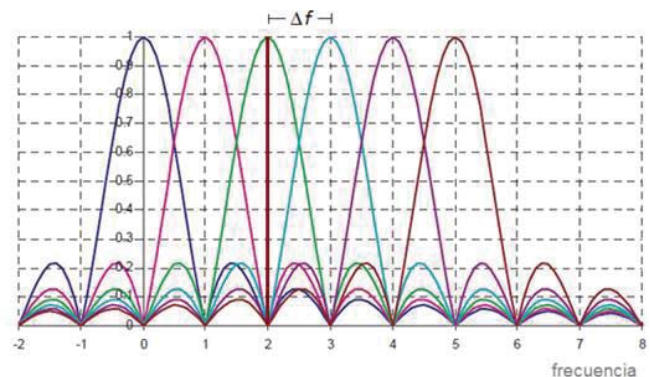


Figura 1. Espectro correspondiente a 6 subportadoras OFDM.

Las ventajas que ofrece OFDM y OFDMA es que al usar múltiples subportadoras de banda estrecha se hace frente al problema de desvanecimientos selectivos en frecuencia. Cada subportadora se somete a su propio desvanecimiento, y si el ancho de banda de cada subportadora es lo suficiente pequeño (mucho menor que el ancho de banda de coherencia), el desvanecimiento selectivo en frecuencia se reduce sustancialmente. Otra ventaja para un sistema multiportadora es que es eficaz en la lucha contra la interferencia entre símbolos (ISI) y el desvanecimiento por trayectorias múltiples. En el dominio del tiempo, los trayectos múltiples conducen a expandir el tiempo de llegada de las señales recibidas debido a múltiples trayectorias de propagación a través de las cuales viajan las señales; al dividir el ancho de banda disponible entre varias portadoras, el tiempo de símbolo se alarga con lo que se minimizará el efecto del interferente anterior [1].

II. TRANSMISOR Y RECEPTOR OFDMA

En la Figura 2 se muestran los bloques que conforman el transmisor y el receptor OFDMA basado en el estándar IEEE 802.16 [1].

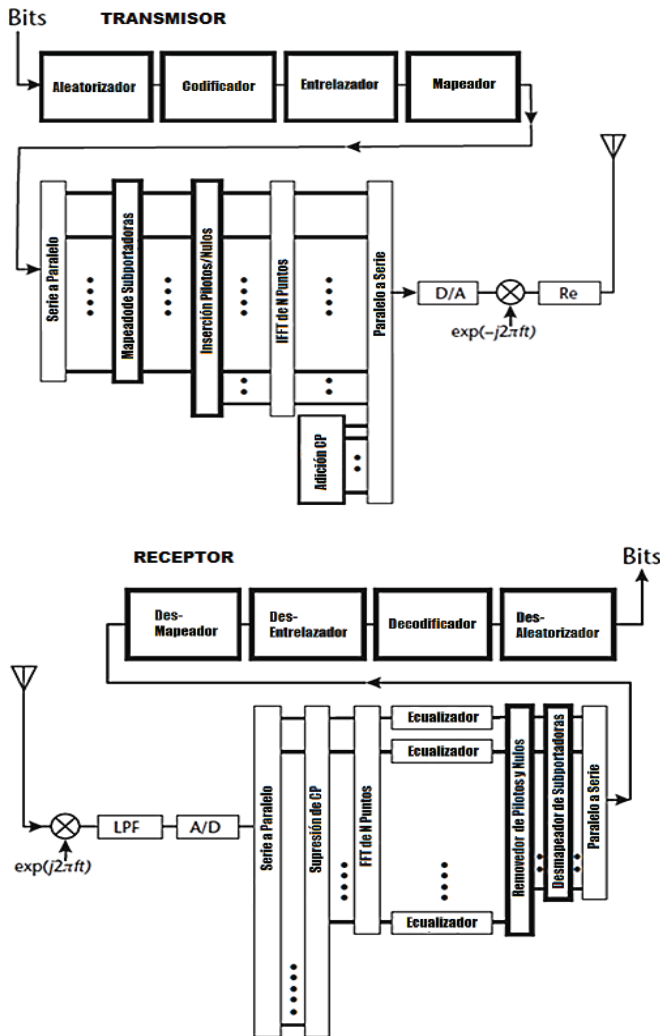


Figura 2. Transmisor y Receptor OFDMA.

A. Aleatorizador/Desaleatorizador

La aleatorización se emplea para minimizar la posibilidad de transmisión de una portadora no modulada y para garantizar un número adecuado de transiciones de bit. Además para distribuir los bits para evitar cadenas largas de ceros y unos [5].

El generador de la secuencia binaria pseudo-aleatoria (PRBS) usado para la aleatorización se muestra en la Figura 3. Los datos entran de forma secuencial en la aleatorización, el valor inicial de los registros será usado para calcular los bits aleatorios, que se combinan mediante una operación XOR con el flujo de bits en serie de cada ráfaga. La secuencia de aleatorización sólo se aplica a los bits de información [6]. Para este bloque en particular su contraparte en el receptor es la misma, donde el valor inicial de los registros debe ser el mismo que en el transmisor para poder recuperar la secuencia original.

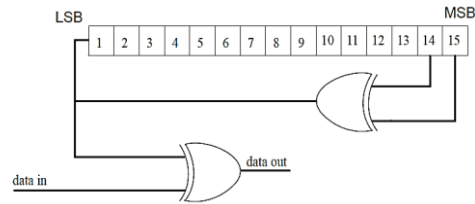


Figura 3. Bloque para la aleatorización de datos.

B. Codificador/Decodificador

La corrección de errores hacia delante o FEC (Forward Error Correction), es un método que permite al receptor corregir datos erróneos sin necesidad de una retransmisión de la información original. Se utiliza para sistemas que operan en tiempo real, donde no se puede esperar a la retransmisión para corregir los datos. La corrección de errores se realiza agregando al mensaje original bits de redundancia. En el estándar IEEE 802.16 se define un FEC conformado por la concatenación de un código Reed-Solomon (externo), y un método de codificación convolucional (interno). Un código convolucional es un código lineal, donde la suma de dos palabras de código, no importa cuales, es también una palabra de código. La implementación de este tipo de código nos da una codificación continua, donde la secuencia de bits codificada depende, además de los bits actuales, de los bits previos. El código trabaja de la siguiente manera: por cada  $k$  bits de información, cuando se codifican se obtienen  $n$  bits, donde  $k/n$  es la tasa de codificación, por ejemplo en una tasa de codificación de  $1/2$  por cada bit de información se agrega un bit de redundancia. Esta codificación se realiza usando un registro de desplazamiento, formado por  $m$  flip-flops y una lógica combinatorial implementada por compuertas XOR, como se ilustra en la Figura 4 [5]:

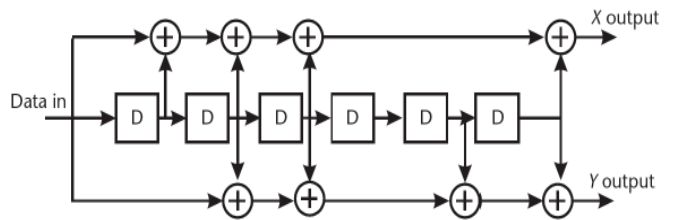


Figura 4. Codificador convolucional de 1/2.

En el receptor el decodificador se encarga de la búsqueda de errores y la corrección de estos. La implementación de este bloque está basada en el algoritmo de Viterbi, cuyo principio es la generación de un algoritmo que encuentre el camino más probable de una secuencia de entrada a partir del camino con menor costo, esto es posible mediante el cálculo de la distancia de Hamming que se obtiene de la comparación entre los bits recibidos y los esperados. Si los bits recibidos son iguales a los esperados se tiene una distancia de cero, si los bits no coinciden se tiene una distancia igual al número de bits diferentes. Para el caso de los códigos convolucionales, los diferentes caminos se expresan mediante el uso de diagramas de Trellis cada camino en el diagrama representa un costo o distancia de Hamming, el algoritmo debe comparar los costos de los caminos y decidir cuál tiene la mayor probabilidad de haber sucedido [7]. En la Figura 5 se muestra el diagrama de Trellis correspondiente al codificador en el transmisor, en donde se tienen 64 estados, que son todas las combinaciones posibles de los valores que pueden tener los flip-flops, además se inicia la decodificación considerando que todos los flip-flops comienzan con un valor de cero, y que siguiendo el algoritmo solo se tienen dos estados posibles. En este diagrama de Trellis se ejemplifica una secuencia de cuatro bits recibidos, en donde un bit no coincide con los esperados.

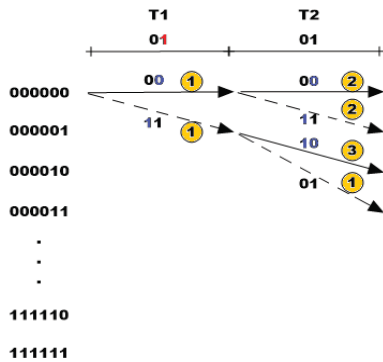


Figura 5. Diagrama de Trellis.

C. Entrelazador/Desentrelazador

El entrelazado se hace para proteger los datos de errores consecutivos durante la transmisión, la idea básica del entrelazado es variar el orden en que se encuentran originalmente los datos a enviar, y que posteriormente serán transmitidos a través de las diferentes subportadoras. De esta forma hablando en términos de OFDM, cada bloque de bits entrelazados se verá afectado por desvanecimientos independientes. Si la mayoría de los bits se reciben de forma correcta, los pocos errores que se produzcan podrán ser corregidos por el decodificador. En lo que se refiere a el proceso de desentrelazado en el receptor, se reorganizan los datos en el orden original. En la Figura 6 se muestra el proceso de entrelazado.

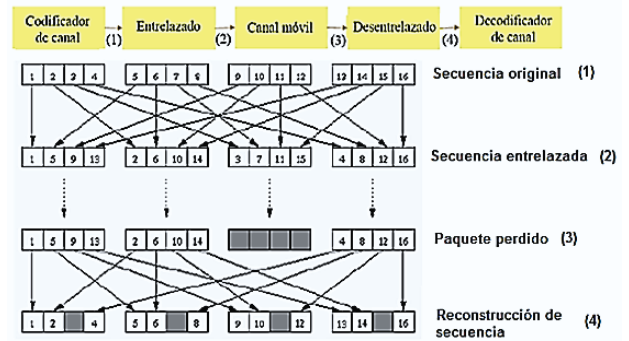


Figura 6. Proceso de entrelazado.

D. Mapeador/Desmapeador

El mapeo se realiza tomando grupos de bits que son asignados a puntos de una constelación. Una magnitud y fase específica representan una determinada combinación de bits. Así una secuencia de bits son mapeados a una secuencia de símbolos de modulación  $S(k)$ , donde  $k$  corresponde al número de símbolo. El número de bits por símbolo depende del tipo de modulación. En el estándar IEEE 802.16 se especifica la modulación BPSK, QPSK, 16QAM y 64QAM. Un símbolo de modulación  $S(k)$  tendrá la forma  $I+jQ$ . La Figura 7 describe el mapeo de bit para una modulación QPSK que es la empleada en este trabajo. El desmapeador en el receptor obtienen los bits de los símbolos modulados, considerando que estos estén dispersos debido a los efectos del canal durante la transmisión.

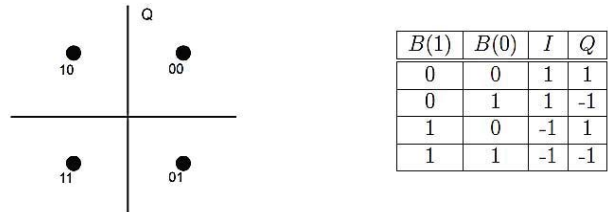


Figura 7. Mapeo utilizando QPSK.

E. Estructura del símbolo OFDM

La estructura del símbolo OFDM está compuesto por subportadoras las cuales pueden ser: de datos; pilotos que son usadas para la estimación de canal; sub-portadoras nulas que son utilizadas como bandas de guarda, y una subportadora de DC. La Figura 8, muestra la estructura de un símbolo OFDM [5]. Esta estructura es considerada en los bloques anteriores a la IFFT, las subportadoras faltantes son agregadas siguiendo el estándar IEEE 802.16.

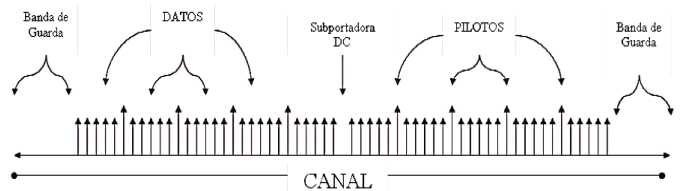


Figura 8. Diagrama de las subportadoras para OFDM.

### F. Transformada Rápida de Fourier/ Transformada Inversa Rápida de Fourier

Tanto en el transmisor como en el receptor de un sistema OFDMA existen un bloque donde se desea cambiar la señal del dominio de la frecuencia al dominio del tiempo y viceversa lo cual conlleva calcular la IFFT y FFT, este bloque es la columna vertebral de este tipo de sistemas ya que gracias a él se realizan los procesos de modulación y demodulación de múltiples símbolos con la certeza de que son ortogonales entre sí. Esta característica de suma importancia es la que permite la transmisión simultánea de múltiples símbolos [3].

### III. ESPECIFICACIONES DEL TRANSMISOR Y RECEPTOR

En la figura 2 se muestra el diagrama del transmisor y el receptor en el cual está basado este trabajo, donde todas las especificaciones fueron tomadas del estándar IEEE 802.16, estas se listan a continuación:

- El número total de puntos para el bloque de la función de la IFFT empleado es de 128, por lo tanto tendremos 128 subportadoras, de las cuales 96 corresponden a los datos, 8 a pilotos, 23 para la banda de guarda y una de DC.
- El tipo de modulación digital usada es QPSK.
- Se usa un codificador convolucional con una tasa de codificación de 1/2.
- El ancho de banda del canal nominal es:  $BW=1.25$  MHz.
- La relación del tiempo de guarda con el tiempo útil del símbolo es:  $G=1/4$ .
- El espacio entre subportadoras está dado por:  $\Delta f=11125$  Hz.
- El tiempo de símbolo útil es:  $T_b = 8.9888 \times 10^{-5}$  s.
- El tiempo de guarda (prefijo cíclico) es:  $T_g = 2.2472 \times 10^{-5}$  s.
- El tiempo del símbolo OFDM es:  $T_s = 1.1236 \times 10^{-4}$  s.

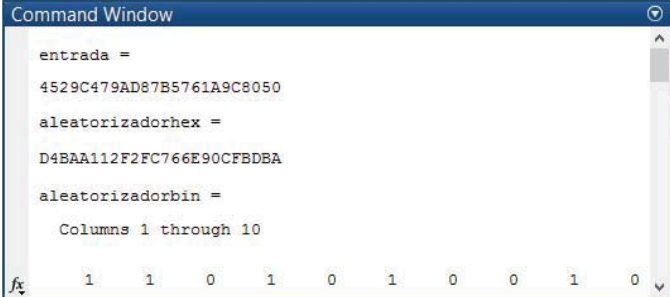
### IV. RESULTADOS

La simulación del sistema OFDMA se realizó en MATLAB tomando como referencia el estándar IEEE 802.16, para posteriormente desarrollar el código en AHDL para su implementación en el FPGA. A continuación se presentan algunos de los resultados obtenidos de las simulaciones tanto en MATLAB como en QUARTUS II de los bloques desarrollados.

#### A. Aleatorizador/Desaleatorizador

En la Figura 9 (a) se presentan los resultados de la simulación en MATLAB, en donde se despliegan los valores de entrada y salida del bloque de la aleatorización. Estos se presentan en hexadecimal para poder visualizar mejor los resultados ya que se tienen 96 bits de entrada y 96 bits de salida, también es mostrada una parte en binario con el objetivo de comparar los resultados con los de QUARTUS II. En la Figura 9 (b) se presenta el resultado de la simulación en QUARTUS II en donde se tiene la señal ENTRADA, la señal ALEA\_TX correspondiente a la salida del aleatorizador, y por

último la señal ALEA\_RX corresponde a la salida del desaleatorizador. Esta señal es igual que la señal de entrada, con lo que se verifica que se está realizando correctamente el proceso inverso a la aleatorización.

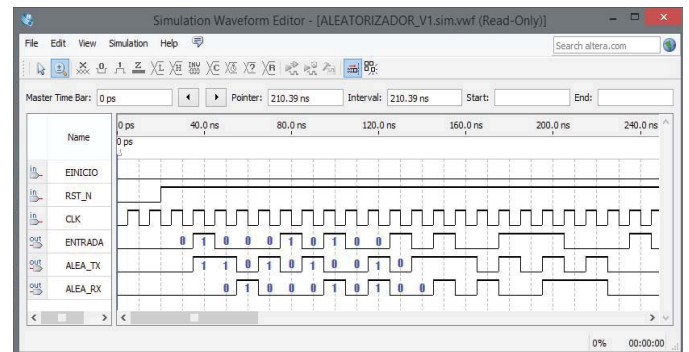


```

Command Window
entrada =
4529C479AD87B5761A9C8050
aleatorizadorhex =
D4BAA112F2FC766E90CFBDBA
aleatorizadorbin =
Columns 1 through 10
1 1 0 1 0 1 0 0 1 0

```

(a) MATLAB.



(b) QUARTUS II.

Figura 9. Simulaciones correspondientes a la aleatorización.

#### B. Codificador/Decodificador

Para comprobar que el decodificador implementado cumpliera con su objetivo, se consideraron dos casos, en el primero las salidas del codificador son introducidas a las entradas del decodificador, en el segundo las salidas del codificador son modificadas, generando algunos errores en las secuencias. Después estos datos son introducidos al decodificador. En la Figura 10 se muestran los resultados obtenidos con MATLAB, en la parte superior se tiene la secuencia de entrada y después las salidas del codificador pero con errores, con el programa implementado se contabilizan los errores encontrados y se despliegan los datos decodificados, en la simulación se puede observar que se corrigieron los 23 errores encontrados, ya que la salida del decodificador corresponde con los introducidos a la entrada del codificador.



```

Command Window
entrada =
4529C479AD87B5761A9C8050
SX_HEX =
DE4E88D5C36262EB8A4C0267
SY_HEX =
57F970D8DA848CF333F7AE0D
conta_error =
23
datos_dec =
4529C479AD87B5761A9C8050
fx >>

```

Figura 10. Decodificación en MATLAB.

En la Figura 11 se muestran los resultados obtenidos en la simulación con QUARTUS II, en donde se tiene la señal SX\_TX y SY\_TX que son las salidas del codificador, seguido se encuentran las señales SX\_RX y SY\_RX generadas con un bloque de prueba, muy semejantes a las salidas del codificador pero con algunas variaciones o errores, que nos sirvieron para evaluar la capacidad del decodificador para corregir los errores, la señal ERROR se muestra en alto cuando un error es detectado por el decodificador, por último se tiene la señal SDECO y la señal SDECO\_SE, donde SDECO\_SE es la señal obtenida del decodificador sin introducir errores, en la simulación se observa que esta corresponde con la señal de entrada, con lo que se verifica que se está realizando la decodificación correctamente, pero además la señal SDECO corresponde a la decodificación introduciendo errores, es importante destacar que el decodificador se probó con error, en donde se obtiene también la secuencia original de entrada, con esto se prueba que el decodificador corrigió los errores.

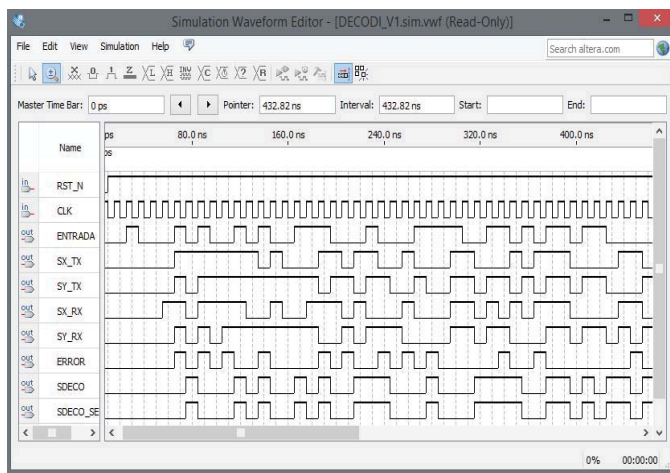


Figura 11. Simulación de la decodificación en QUARTUS II.

### C. Entrelazador/Desentrelazador

En la Figura 12 se presentan los resultados obtenidos en la simulación con QUARTUS II del proceso de entrelazado, para el diseño implementado se tienen dos entradas al bloque de entrelazado, ya que el bloque anterior es el codificador y este tiene dos salidas, entonces las señales a entrelazar son ENTRADA1 y ENTRADA2, las señales ENTB1\_TX y ENTB2\_TX son las salidas de este bloque, el tiempo requerido para realizar el entrelazado es de un  $1\mu s$ , puesto que es necesario primero almacenar todas las secuencias de entrada. Posteriormente los datos son mostrados en el orden calculado por las expresiones que aparecen en el estándar IEEE 802.16. En el receptor se realiza un proceso similar con el bloque de desentrelazado pero las entradas a este bloque provienen del desmapeador de símbolos. En la prueba realizada la salida del entrelazador es introducida al bloque del desentrelazador. En la simulación se observa que se obtiene nuevamente la secuencia original de entrada al entrelazador, con lo que se prueba que el proceso de entrelazo se realiza correctamente.

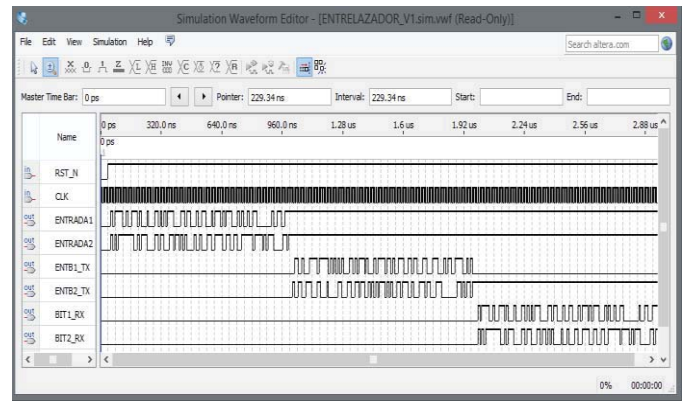


Figura 12. Simulación del proceso de entrelazado en QUARTUS II.

### D. Mapeador/Desmapeador

En la figura 13 se muestran los resultados obtenidos de la simulación en MATLAB del mapeo de símbolos, la gráfica del lado izquierdo muestra los símbolos enviados y la gráfica del lado derecho muestra los símbolos recibidos después de pasar por un canal con ruido.

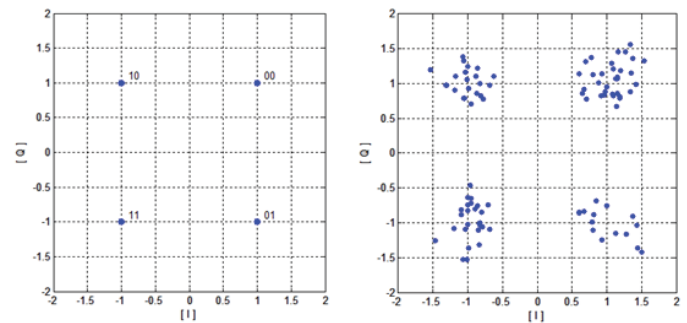


Figura 13. Mapeo en MATLAB.

En la Figura 14 se muestran los resultados obtenidos en la simulación con QUARTUS II, a la entrada del mapeador de símbolos se tienen las señales ENT1 y ENT2, y a la salida de este bloque se tiene SREAL y SIMAG, formados por 20 bits, estos dos números sirven para representar los valores de los ejes de la constelación QPSK, en el receptor se realiza el proceso inverso, según el valor de SREAL y SIMAG se obtiene un par de bits.

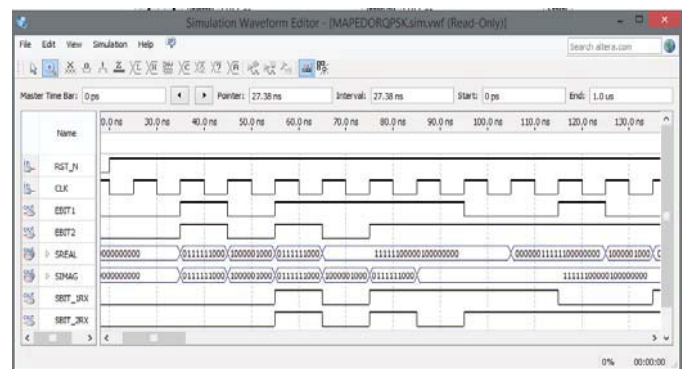


Figura 14. Simulación del mapeo de símbolos en QUARTUS II.

### E. Transformada Rápida de Fourier/Transformada Inversa Rápida de Fourier

El bloque de la IFFT y FFT ya fue desarrollado en otro trabajo, este bloque ya fue implementado en un FPGA Cyclon III, en el diseño de este bloque se consideró reducir el mayor número de operaciones para el cálculo de la IFFT/FFT y ocupar el menor número de elementos lógicos posible. La IFFT y FFT se puede realizar con el mismo bloque, y el tiempo necesario para realizar la transformada es de 34 $\mu$ s [8].

### V. CONCLUSIONES

Las simulaciones en MATLAB ayudaron a comprender mejor el sistema OFDMA, nos permitió realizar un estudio más a fondo de cada uno de los bloques que conforman todo el sistema, probar diferentes entradas en los bloques y analizar el sistema considerando ruido durante la transmisión, además de probar distintos diseños y algoritmos antes de ser implementados en el FPGA.

En las simulaciones realizadas en QUARTUS II se mostró el correcto funcionamiento de los bloques del transmisor y receptor. Un bloque a resaltar es el decodificador en el receptor el cual es capaz de corregir errores, algo importante en este tipo de sistemas.

Algo que se debe considerar en un sistema OFDMA es el tiempo del símbolo a transmitir, para nuestro diseño es de 112.3  $\mu$ s, lo que indica que todo el procesamiento tanto en el transmisor como en el receptor tiene que ser menor a ese tiempo, tomando en cuenta que la IFFT o FFT tarda 34  $\mu$ s y

sumando el tiempo en procesar los datos en los demás bloques se llega a un tiempo menor de 40  $\mu$ s, con lo que se cumple con este parámetro.

Como trabajo a futuro restan las pruebas con todos los bloques en conjunto y la verificación del procesamiento de los datos de todo el sistema.

### REFERENCIAS

- [1] Samuel C. Yang, OFDMA System Analysis and Desing. Artech House,2010.
- [2] LaSorte, N., W. Barnes, J., & H. Refai, H. Multiplexing, The History of Orthogonal Frequency Division. Communications Magazine, IEEE, 26-35,2009.
- [3] J. Vergara González, Simulación de un Esquema de Modulación/Demodulación OFDM Utilizando un Modelo de Canal Multitrayectoria. Guayaquil, Ecuador,2008.
- [4] R. Agusti, F. Bernardo, F. Casadevall, R. Ferrús, J. Pérez y O. Sallent, LTE: Nuevas tendencias en comunicaciones móviles, Fundación Vodafone España, 2010.
- [5] IEEE Standard 802.16," IEEE standard for Local and metropolitan area networks, Part 16: Air Interface for Broadband Wireless Access Systems". IEEE Microwave Theory and Techniques Society, 2009.
- [6] A. Gutiérrez, Diseño del nivel físico de un radio OFDM para comunicación digital con una plataforma SDR, Enero 2014.
- [7] H. Díaz, O. Hurtado, Implementación en software de un decodificador de Viterbi para aplicaciones DVB-S, Instituto Politécnico Nacional, 2013.
- [8] F. Mora Hernández, J. Castañeda Camacho, S. Vergara Limón, E. Ríos Silva, J. E. M. Gutiérrez Arias, Implementación de la Transformada Discreta de Fourier IFFT/FFT en un sistema embebido para automatizar un sistema de comunicación OFDMA, CIINDET, 2015.

# Implementación de una Red Neuronal Artificial para la detección de personas en un Hardware Reconfigurable

Guerrero Ramírez, Carlos Francisco  
Maestría en Sistemas  
Computacionales  
ITLP  
La Paz, B.C.S, México  
c.guerreros@outlook.com

Santillán Méndez, Israel Marcos  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S, México  
santilis@gmail.com

Martínez Díaz, Saúl  
División de Estudios de  
Posgrado e Investigación  
ITLP  
La Paz, B.C.S, México  
saulmd@itlp.edu.mx

**Resumen**—La investigación descrita en este artículo explora la factibilidad de implementar un sistema de reconocimiento de objetos, en una secuencia de imágenes bidimensionales diferenciando entre humano y no-humano, en un dispositivo semiconductor reconfigurable. Uno de los objetivos de realizar esta implementación es comparar el tiempo de procesamiento entre el dispositivo semiconductor reconfigurable y una PC convencional. Para hacer frente a este reto se ha empezado por implementar una Red Neuronal Artificial (RNA) Feedforward en una PC convencional, la cual ha sido entrenada con el algoritmo Backpropagation. Se ha utilizado el método de momentos estándar, para la extracción de características de las imágenes, para realizar el entrenamiento de la RNA. El sistema implementado puede reconocer perfiles de humanos con un éxito de un 90%.

**Palabras clave**—Redes Neuronales; Backpropagation; Momentos estándar; FPGA.

## I. INTRODUCCIÓN

### 1. Redes Neuronales Artificiales.

Las Redes Neuronales Artificiales (RNA) son un sistema de procesamiento de información que ayuda a simular la estructura del cerebro humano y sus funciones. Las RNA han sido utilizadas para resolver problemas en donde los métodos convencionales no han sido exitosos, tales como el procesamiento de imágenes, el reconocimiento y clasificación de patrones, etcétera [1].

En 1943, el neurobiólogo McCulloch y el matemático Pitts propusieron la primera red neuronal artificial, conocida como el Modelo MP [2]. Así comenzaba la era de investigación de las RNAs. Después, en 1958, un dispositivo de procesamiento en paralelo basado en la biología humana, conocido como el Perceptrón sería inventado por Frank Rosenblatt [3]. El perceptrón era capaz de desarrollar operaciones computacionales lógicas sencillas (específicamente aquellas que requirieran únicamente la diferenciación entre patrones de entrada linealmente separables. Más tarde, en 1969, Marvin Minsky y Seymour Papert, publicarían un libro titulado: El Perceptrón: Una Introducción a la Geometría Computacional,

en el cual señalaron que incluso la sencilla tarea de predecir la salida de una compuerta XOR estaba más allá de las capacidades del perceptrón [4].

Pasarían casi 20 años para que en 1986, una capa oculta fuera añadida entre las capas de entrada y las capas de salida para el entrenamiento de las redes neuronales artificiales y de esta manera realizar la tarea de clasificar patrones más complejos. Esta mejora al modelo del Perceptrón es llamada RNA Backpropagation (BPNN, por sus siglas en inglés), la cual fue presentada por Rumelhart y McClelland [5].

### 2. FPGA

Los FPGA son dispositivos semiconductores reconfigurables que fueron desarrollados a mediados de 1980. Una computadora personal puede ser usada para diseñar un circuito digital el cual después es compilado en un archivo de programación especial, el cual se ejecutará en el circuito una vez que haya sido descargado al FPGA.

Los FPGAs disponibles de los fabricantes, actualmente consisten de miles de elementos lógicos reconfigurables. Cada uno de estos elementos lógicos contienen Flip-Flops, Look-up Tables, compuertas NAND, y compuertas NOR. El cableado que conecta estos componentes, así como el cableado que conecta los elementos lógicos vecinos, son eléctricamente reconfigurables. Por lo tanto los elementos pueden ser cableados para crear un diseño, y después borrados y cableados de diferente manera para crear un nuevo diseño diferente [6] [7].

## II. JUSTIFICACIÓN

Las RNA son una descripción abstracta del cerebro humano. Estas al ser un modelo matemático, pueden ser implementadas en circuitos integrados o simuladores usando programas computacionales. No obstante, el paralelismo inherente embebido en la dinámica de las RNA solo puede ser plenamente utilizado en una implementación de hardware.

Las computadoras de arquitectura Von Neumann son bien conocidas por realizar simulaciones de RNAs. Sin embargo, la

velocidad de este tipo de simulaciones es limitada cuando el tamaño de las RNAs llegan a ser grandes. Adicional a esto, cabe mencionar que las simulaciones en software son ejecutadas secuencialmente.

### III. METODOLOGÍA

La primera fase de esta investigación ha consistido en obtener una base de datos con imágenes de personas. La base de datos de CASIA [8] que se ha seleccionado cuenta con poco mas de 19,000 imágenes que han sido previamente segmentadas y muestran solo la silueta de una persona. La base de datos posee imágenes binarias de siluetas de personas en diferentes posiciones y tamaños. Hasta el momento solo se han seleccionado 75 imágenes, las cuales tienen un tamaño de 240x352, una de ellas se puede observar en la figura 1.



Figura 1. Imagen binaria de la silueta de una persona.

Una vez que hemos obtenido la base de datos de imágenes con objetos que hemos etiquetado como humano, también ha sido necesario obtener imágenes con objetos que etiquetaremos como no-humano. Una vez que hubo quedado lista la base de datos de imágenes, lo que sigue es realizar la extracción de las características.

Generalmente el reconocimiento de objetos se lleva a cabo utilizando ciertos rasgos característicos del objeto buscado. Para reconocer objetos del mundo real, es deseable que los rasgos seleccionados sean invariantes a rotación y escalamiento. Una técnica ampliamente utilizada para este propósito se basa en el cálculo de momentos estadísticos. Para imágenes digitales el momento estadístico (1) de orden (p+q) puede calcularse por:

$$(1) \quad m_{pq} = \sum_x \sum_y x^p y^q f(x, y),$$

Donde p,q= 0, 1, 2,...

Dado que p y q pueden tomar una cantidad infinita de valores, es posible generar una cantidad infinita de momentos. Este conjunto determina de manera univoca cada función f(x,y) y viceversa. Para lograr invariancia a traslación se definen los momentos centrales (2):

$$(2) \quad \mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y)$$

$$\text{Donde } \bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}$$

Hu define siete momentos normalizados (3) (4) (5) (6) (7) (8) (9) [9], los cuales son invariantes a rotación, traslación y escalamiento:

$$(3) \quad \phi_1 = \eta_{20} + \eta_{02}$$

$$(4) \quad \phi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2$$

$$(5) \quad \phi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2$$

$$(6) \quad \phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2$$

$$(7) \quad \phi_5 = (\eta_{30} + 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \dots \\ + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 \dots \\ - (\eta_{21} + \eta_{03})^2]$$

$$(8) \quad \phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \dots \\ + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03})$$

$$(9) \quad \phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 \dots \\ - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{03})(\eta_{21} - \eta_{03}) \dots \\ [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]$$

$$\text{Donde } \eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^\gamma} \text{ con } \gamma = \frac{p+q}{2} + 1$$

Para p+q = 2,3, 4...

Los momentos pueden ser usados para derivar un conciso, altamente descriptivo vector característico de una imagen bidimensional. Fueron computados una total de 264 vectores característicos de los cuales 86 fueron etiquetados como objeto humano, y se les identifico con el valor 1. Los restantes 178 valores fueron etiquetados como objetos no-humanos, y se les identifico con el valor 0.

Una vez definida también la manera en la cual se van a extraer los valores característicos, fue necesario definir la RNA que se utilizaría. Una RNA Feedforward multicapa, como se ilustra en la figura 2, está compuesta por una entrada, una salida, y una o más capas ocultas. Con la excepción de la capa de entrada, todas las capas computan su salida, usando una formula de salida ponderada, un bias opcional y una función de activación. Para lograr un aprendizaje supervisado se utiliza el algoritmo Backpropagation [10]. Rao y Rao [10] desarrollaron la red neuronal en la que está basada este sistema. El algoritmo



Backpropagation se uso para clasificar la información de los vectores obtenidos mediante el método de momentos invariantes en clases predeterminadas definidas previamente.

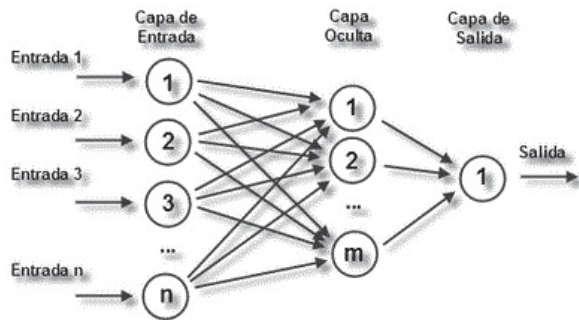


Figura 2. Red Neuronal Artificial Feedforward

IV. PROCEDIMIENTO

La figura 3 muestra el procedimiento de cómo el sistema de reconocimiento utiliza los vectores característicos obtenidos mediante el método de momentos invariantes [9] para el reconocimiento de objetos. El procedimiento está influenciado en Nicolaou, C. A. et al. [11]. Durante la fase de entrenamiento se ha introducido el 70% de nuestro conjunto de vectores de momentos al sistema de reconocimiento de patrones. Entonces la BPNN es entrenada para reconocer el etiquetado de imágenes previamente realizado. Durante la fase de prueba, el 15% restante de nuestro conjunto de vectores de momentos “desconocido” es clasificado, y la salida es comparada con la etiqueta previamente definida del vector característico para verificar la coincidencia. De esta manera se puede realizar el cálculo del porcentaje de rendimiento de nuestro sistema, identificando los aciertos positivos, los falsos positivos, y los falsos negativos.

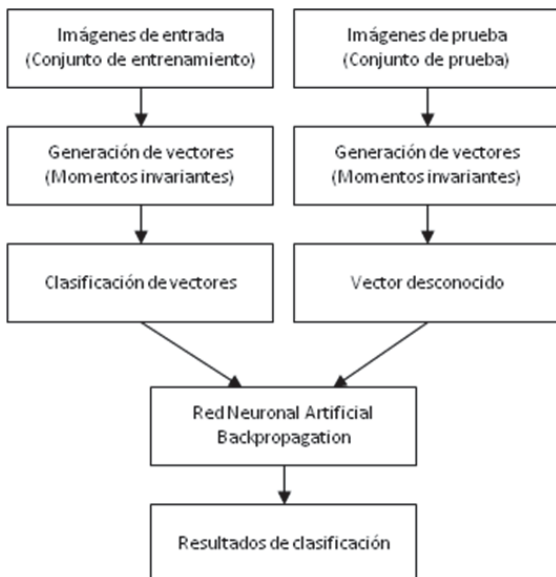


Figura 3. Reconocimiento de objetos basado en Momentos invariantes

V. RESULTADOS

El porcentaje de error del sistema de reconocimiento es del 5% al 10%. A continuación se presentan los resultados obtenidos.

	Numero	Porcentaje
Aciertos Humanos	54	29.35
Aciertos No-Humanos	120	65.22
Falsos Positivos	6	3.26
Falsos Negativos	4	2.17
Total	184	100

Tabla 1. Resultados del entrenamiento

	Numero	Porcentaje
Aciertos Humanos	13	32.5
Aciertos No-Humanos	23	57.5
Falsos Positivos	0	0
Falsos Negativos	4	10
Total	40	100

Tabla 2. Resultados de la validación

	Numero	Porcentaje
Aciertos Humanos	10	25
Aciertos No-Humanos	28	70
Falsos Positivos	1	2.5
Falsos Negativos	1	2.5
Total	40	100

Tabla 3. Resultados de prueba

	Numero	Porcentaje
Aciertos Humanos	77	29.2
Aciertos No-Humanos	171	64.8
Falsos Positivos	7	2.6
Falsos Negativos	9	3.4
Total	264	100

Tabla 4. Resultados totales

Para tener un punto de comparación en cuanto a los resultados obtenidos en esta sistema, cabe mencionar que en la investigación realizada por Nicolaou, C. A. et al. [11], sus resultados de aciertos humanos positivos son de un 87.5% a un 92.5%. Y el porcentaje para los aciertos no-humanos están entre un 87.5% a un 96.5%

VI. CONCLUSIONES Y TRABAJO FUTURO

Al momento de redactar este artículo, aun continua el proceso de la implementación de la RNA en el FPGA.

El reconocimiento de personas es un área de la visión artificial que permite se realicen importantes aportaciones en diversas áreas de la sociedad como la medicina, la seguridad ciudadana, entre otras.

Dada la cantidad de información que se tiene que procesar es importante el uso de tecnologías que permitan sacar el máximo rendimiento a esos algoritmos que permite la paralelización de tareas. El uso de FPGA nos da esta capacidad, así como la posibilidad de implementar sistemas autónomos que demanden poco consumo de recursos.

En cuanto al área de la visión artificial es un área en constante evolución. Cada vez se descubren nuevas técnicas para lograr una mejor obtención de valores característicos de los objetos del mundo real capturados a través de una secuencia de imágenes bidimensionales. Dicho esto, está en los planes implementar un Sistema de Visión Estéreo (SVE) que permita realizar la segmentación del objeto de interés de una mejor manera, y además proporcione información adicional, propia de un SVE.

Adicional al SVE, se busca incrementar el número de imágenes de los cuales se extraen los valores característicos mediante el método de momentos.

#### REFERENCIAS

- [1] M. Ananda Ro and J. Srinvas. *Neural Networks Algorithms and Applications*, Alpha Science International Ltd., Part III, pp. 157, 2003.
- [2] McCulloch W S and Pitts W., "A logical calculus of the ideas immanent in nervous activity", *Bulletin of Mathematical Biophysics*, (5), pp. 115-133, 1943.
- [3] F. Rosenblat, "The Perception: A probabilistic model for information storage and organization in the brain". *Psychological Review* Vol. 62, No. 3, pp. 386-408, 1958.
- [4] Minsky, M. Seymour, P. "Perceptrons. An introduction to computational geometry: Cambridge: MIT Press, 1969.
- [5] J. L. McClelland and D. E. Rumelhard, "Exploration in parallel distributed processing", *A handbook of Models, Programs, and Exercises*. Cambridge: MIT Press, 1986.
- [6] Zh. Wang and E. Zhu, *VLSI design*, Beijing: Publishing House of Electronic Industry, pp. 75-92, 2005.
- [7] R. Zeidman, *Designing with FPGAs and CPLDs*. Kansas: CMP Books, 2002.
- [8] Liang, Hongyu. "Casia GAIT Database: Dataset A", de <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>
- [9] M. Hu. "Visual Pattern Recognition by Moment Invariants". *IRE Trans Info. Theory*, 1962, vol. 8, pp. 179-187.
- [10] V.B. Rao and H.V. Rao, *Neural Networks and Fuzzy Logic*, MIT Press, second edition, 1995.
- [11] Nicolaou, C.A., Jr. Egbert A.L., R.C. Lacher, and S.I. Bassett. "Human Shape Recognition Using the Method of Moments and Artificial Neural Networks." In *International Joint Conference on Neural Networks, 1999. IJCNN '99*, 5:3147-51 vols.5, 1999.

# Implementación del algoritmo de ICA en un DSP para la atenuación de ruido en llamadas telefónicas

Luz Noé Oliva Moreno  
Instituto Politécnico Nacional  
IPN - UPIIH  
México, Estado de Hidalgo  
loliva@ipn.mx

Chávez Barrera Rosa Graciela  
Pasante Ing. Sistemas Computacionales  
ESCOM-IPN  
México, Ciudad de México  
zso.boo@gmail.com

**Abstract**—Este trabajo presenta la implementación de un algoritmo de Análisis de Componentes Independientes (ICA) en un Procesador de Señales Digitales (DSP) de Texas Instruments que atenúa el ruido en las llamadas telefónicas. El algoritmo utilizado es el H-J, el cual fue desarrollado por Christian Jutten y Jeanny Héroult [1] y mediante algunas reglas en las condiciones de operación evita la saturación de los pesos. Los resultados arrojaron una atenuación del ruido de aproximadamente 35 dB.

**Keywords**—Procesador de Señales Digitales, Análisis de Componentes Independientes, Cancelación de Ruido Adaptable, manos libres, Procesamiento de señales.

## I. INTRODUCCIÓN

El procesamiento de señales es utilizado en diferentes áreas y ha sido objeto de estudio durante muchos años. Un problema clásico en el procesamiento de señales es la cancelación del ruido [1], dicho problema ha sido objeto de estudio por diversos investigadores [2], pero aún no se cuenta con una solución perfecta, sin embargo, se han obtenido niveles aceptables de atenuación y se siguen presentado nuevas investigaciones con mejoras del mismo.

Una de las líneas de investigación para la atenuación del ruido se llama Análisis de Componentes Independientes (ICA por sus siglas en inglés), la cual se basa en las propiedades estadísticas de las señales, debido a que esta técnica asume que las fuentes de señales son independientes.

Christian Jutten y Jeanny Héroult [3] desarrollaron un algoritmo llamado H-J, el cual consiste una red neuronal ICA que obtiene las fuentes independientes de una mezcla. Asumiendo que la voz y el ruido son señales independientes, la atenuación del ruido se obtiene una vez que los pesos de la red neuronal convergen. Lo anterior tiene diversas aplicaciones, pero el artículo se centra en implementar un sistema de atenuación del ruido en llamadas telefónicas.

Para la obtención de los resultados del algoritmo H-J, se implementó un sistema en Matlab, el cual contiene dos etapas, la primera modela las mezclas del ruido y la voz, simulando el ambiente cuando se realiza la llamada y la segunda ejecuta el algoritmo.

Uno de los dispositivos diseñados para realizar operaciones de punto flotante en señales son los Procesadores de Señales Digitales o DSP. Se utilizó el DSP TMS320F28335 de la

empresa Texas Instruments para la implementación del algoritmo.

## II. ALGORITMO DE ANÁLISIS DE COMPONENTES INDEPENDIENTES PARA LA CANCELACIÓN DE RUIDO ADAPTABLE

El algoritmo ICA es utilizado en el problema de separación ciega de señales (BSS por sus siglas en inglés) [4]. El problema BSS se refiere a una técnica de separación de señales, sin el conocimiento de alguna de las señales, solo la característica de la independencia estadística [5]. El diagrama a bloques de ICA se muestra en la figura 1:

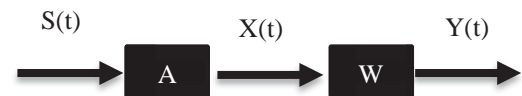


Figura 1. Diagrama a bloques de ICA.

En donde  $S(t)$ , es un vector que representa las señales originales,  $A$  es la matriz de mezcla que se genera en el ambiente donde se propagan dichas señales, y  $X(t)$  es el vector de las señales recibidas por los sensores y corresponden a las mezclas observadas,  $W$  es la matriz de pesos que realiza la transformación inversa de la mezcla [6] y  $Y(t)$  es el vector de las salidas estimadas de las señales originales. Las señales mezcladas se expresan de la siguiente forma:

$$X(t) = AS(t) \quad (1)$$

El vector  $Y(t)$  se obtiene a partir de la transformación  $W$  de las señales  $X(t)$  y se expresa como:

$$Y(t) = WX(t) = WAS(t) \quad (2)$$

Donde  $W$  es la matriz de pesos sinápticos de la red neuronal que realiza una transformación lineal para obtener las fuentes independientes. La solución del problema de BSS consiste en

encontrar una matriz de transformación inversa de la matriz de mezcla A, a partir de las señales S(t).

Para encontrar la solución de la separación de señales originales se utilizó la arquitectura de la red neuronal y su regla de aprendizaje [7] que fue desarrollado por Christian Jutten y Jeanny Héroult. Su regla de aprendizaje se basa en el método de descenso del gradiente y actualiza los pesos sinápticos de la red neuronal donde el ajuste se realiza en forma continua de acuerdo a la siguiente expresión, conocida como la regla de Hebb o regla de adaptación, dada por:

$$\frac{dc_{12}(t)}{dt} = n * f[s_1(t)] * g[s_2(t)] \quad (3)$$

Dónde: *n* es la razón de aprendizaje.  
*f* y *g* son funciones no-líneales

La actualización de los pesos se puede expresar de forma recursiva como:

$$c_{12}(t+1) = c_{12} + n * f[s_1(t)] * g[s_2(t)] \quad (4)$$

Dónde: *c(t+1)* es el peso sináptico siguiente.  
*c* es el peso sináptico actual.

Siendo la ecuación (4) la regla de aprendizaje de la red neuronal. Para lograr una independencia estadística, Jutten y Héroult propusieron el uso de dos funciones no lineales e impares, preservando la condición de media cero.

Experimentalmente fueron determinadas las siguientes restricciones en *f* y *g* asegurando una convergencia rápida y estable de los pesos.

- Para una convergencia estable, *f* debe tener una curvatura positiva y *g* una curvatura negativa.
- Para una convergencia rápida, *f* y *g* deben ser ortogonales en el origen.

Las funciones *f* y *g* propuestas por Jutten y Héroult fueron:

- $f(x) = \tanh(x)$
- $g(x) = x^3$

Y estas funciones son las utilizadas durante el aprendizaje de la red neuronal.

Para que los pesos pudieran converger, se tuvo que realizar una adaptación en cuanto a la amplitud de las señales. Esta adaptación se realizó por software, dividiendo entre 0x7fff todos los valores que ingresaban del Códec. Este valor se eligió porque es el valor máximo que se puede muestrear en 16 bits. De esta forma se evita una saturación de los datos, ya que los pesos se mantienen estables sin que tiendan a infinito.

Para comprobar el desempeño del algoritmo, se modeló un sistema en Matlab [8]. En una de las pruebas realizadas se utilizaron señales sinusoidales [9,10], arrojando los siguientes valores para el cálculo de pesos sinápticos y atenuación en decibels.

Pesos originales	Pesos calculados por el algoritmo
$\begin{pmatrix} 1 & 0.4 \\ 0.6 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0.407 \\ 0.597 & 1 \end{pmatrix}$

**Tabla 1.** Valores de los pesos después de converger el algoritmo

Los decibels atenuados [11] para la señal voz se calcularon con la siguiente fórmula:

$$20 \log\left(\frac{mV \text{ voz}}{mV \text{ ruido}}\right)$$

$$20 \log\left(\frac{0.407}{0.001151}\right)$$

$$20 \log(353.605560382)$$

$$50.9703817119 \text{ dB}$$

### III. IMPLEMENTACIÓN EN DSP

El algoritmo H-J fue implementado en un DSP TMS320F28335, utilizando una tarjeta de desarrollo de Texas Instruments, el cual consta físicamente de dos partes principales, la Control card, que es una tarjeta de 100 pines donde está el DSP y la Docking station que cuenta con diversos puertos periféricos y un CODEC de audio para el muestreo de la señal, además posee conectores jacks de entrada y salida, para ingresar y reproducir audio.

Texas Instruments provee, junto con el hardware de sus microcontroladores y DSPs, un ambiente de desarrollo específicamente diseñado para programar sus dispositivos llamado Code Composer Studio (CCS); el algoritmo se programó usando la versión 4.4 de dicho software.

La comunicación entre el DSP y el CCS se mantiene mediante el módulo emulador XDS100, el cual provee acceso JTAG al DSP desde un puerto USB en la PC. Dicha comunicación puede mantener acceso en tiempo real a memoria y registros de control en el DSP, además de poder contar ciclos de reloj entre dos segmentos de código, graficar segmentos de memoria, ver código desensamblado, entre otras características.



**Figura 2.** TMS320F28335 Experimenter Kit.

Dentro de esta tarjeta también se utilizó el códec que está conectado al DSP.

El códec de audio TMLVAIC23B, tiene las siguientes características:

- Frecuencias de hasta 192KHz.
- Dos entradas de audio (de línea y de micrófono).
- Dos salidas de audio (audífono y línea).
- Comunicación SPI.

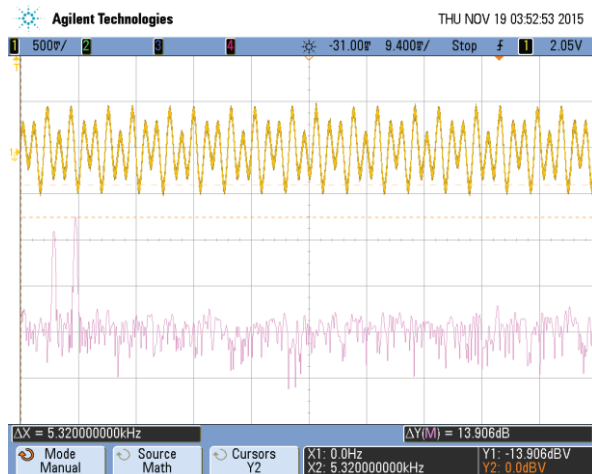
#### IV. RESULTADOS

Para la prueba de señales sinusoidales se utilizaron señales con diferentes frecuencias, creadas por medio de un generador de funciones, y las mezclas se obtuvieron por medio de un sumador analógico.

Para las pruebas con señales de audio reales se realizaron diferentes grabaciones por medio de micrófonos en formato WAV y se crearon las mezclas para ser ingresadas al algoritmo.

#### Señales sinusoidales

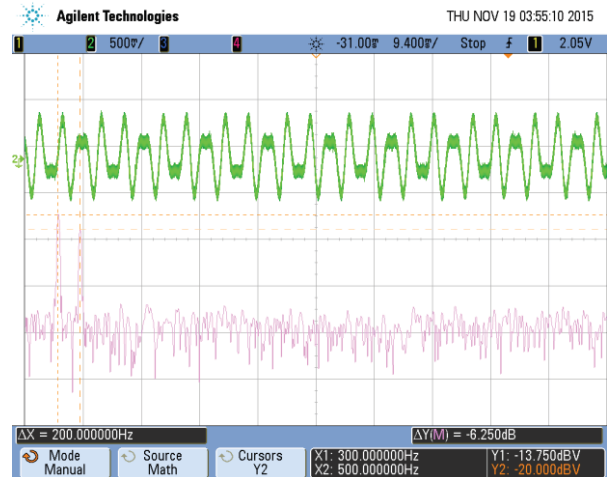
Una de las pruebas que se realizaron fueron con dos señales, una señal seno con frecuencia de 300Hz y señal seno con frecuencia de 500Hz.



**Figura 3.** Mezcla 1.

En la figura 3, se muestra la primera mezcla, con su transformada de Fourier, esta, muestra las componentes de las

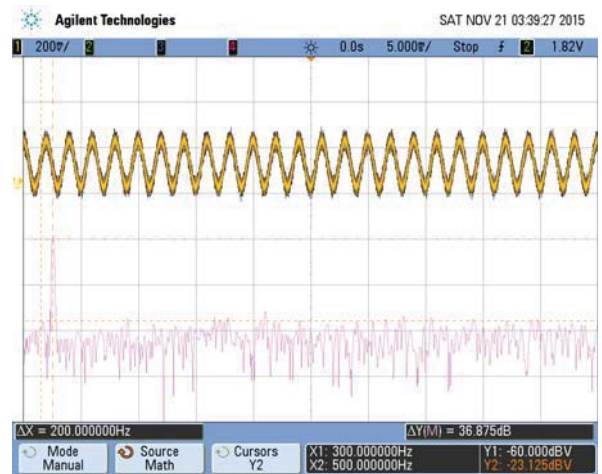
dos señales en la mezcla. La diferencia en decibeles de la amplitud de las señales es de 6.250 dB.



**Figura 4.** Mezcla 2.

En la figura 4, se muestra la segunda mezcla, con su transformada de Fourier, en esta última, se observa la existencia de las dos señales en la mezcla. La diferencia en decibeles de la amplitud de las señales es de 6.250 dB.

Una vez que se tienen ambas mezclas, éstas ingresan al DSP para el cálculo del algoritmo, obteniendo los resultados en la figura 5 y 6:



**Figura 5.** Señal (1) original estimada por el algoritmo.

En la figura anterior, se tiene la salida donde se observa que la amplitud de la señal seno con frecuencia de 300Hz, predomina por encima de la amplitud de la otra frecuencia. La diferencia en decibeles es de 36.875 dB aproximadamente.



Figura 6. Señal (2) original estimada por el algoritmo.

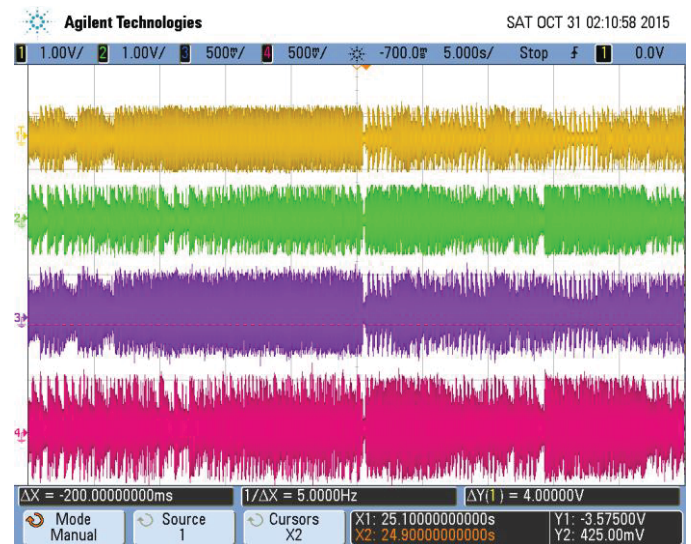


Figura 7. Prueba señales reales.

En la figura 6.19 se tiene la salida del algoritmo y se observa que la frecuencia de 500Hz predomina sobre las otras frecuencias. La diferencia en decibeles es de 45.625 dB aproximadamente.

En cuanto a los pesos sinápticos se analizó un caso de la red neuronal programada en el DSP. Los valores se muestran en la siguiente tabla:

Pesos utilizados	Pesos calculados por el algoritmo
$\begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$	$\begin{pmatrix} 1 & 0.4938082 \\ 0.5070135 & 1 \end{pmatrix}$

Tabla 1. Valores de los pesos utilizados y calculados por el algoritmo.

Otras pruebas que se realizaron, arrojaron los resultados obtenidos de la siguiente tabla:

Señal 1	Señal 2	Relación señal a ruido salida (dB)	Relación señal a ruido salida (dB)	w11	w12
sin 300	sin 500	36.875	45.625	0.49380	0.50701
cos 250	cos 1100	43.750	43.750	0.68940	0.40587
cuad 250	sin 950	38.750	33.125	0.49287	0.50708
trang 500	cuad 800	35.625	39.375	0.67403	0.42949
cos 870	cuad 400	48.125	36.875	0.39434	0.50794

Tabla 2. Resultados obtenidos en las pruebas.

En la figura 7 se muestran las señales originales reales que se obtuvieron con el algoritmo:

La señal 1 y señal 2 son las señales originales, las salidas del algoritmo se muestran en la señal 3 y señal 4 de la figura anterior. Se puede observar que la señal 1 y señal 3 son similares, al igual que la señal 2 y señal 4.

## V. CONCLUSIONES

La implementación del algoritmo en un DSP obtuvo la atenuación del ruido en las llamadas, el cual utilizo un algoritmo de Análisis de Componentes Independientes (ICA) en un Procesador de Señales Digitales (DSP) de Texas Instruments operando en tiempo real. Para evitar la saturación de los pesos se programó algunas reglas en las condiciones de operación. Los resultados arrojaron una atenuación del ruido de aproximadamente 35 dB y se puede apreciar de forma auditiva dichos resultados.

## REFERENCIAS

- [1] W. P. Robins, "Phase Noise in Signal Sources: Theory and Applications". IET, 1984, Technology and Engineerin, 321 pages.
- [2] E.Sorias, "Tratamiento digital de señales: problemas y ejercicios resueltos". Pearson Educación, S.A, Madrid, 2003, pp 400.
- [3] A. Jutten and Herault., "Blind separation of source, Part II: Problems statement". Signal processing, Vol. 24. New York: Elsevier Science, 1991, pp. 11-20.
- [4] A. Hyvärinen, J. Karhunen, E. Oja., "Independent Component Analysis". 3a ed. United States of America: Wiley, 2001.
- [5] A. Papoulis, "Probability, Random Variables and Stochastic Processes". 4th edition, McGraw Hill, 2002.
- [6] S. Laughlin, Z. Naturforsch, "A simple coding procedure enhances a neuron's information capacity", pp. 36 (9-10):910-912,1981.

- [7] Bell, A. J. And Sejnowski, T.J., "An Information-Maximization Approach to Blind Separation and Blind Deconvolution". *Neural Computation*, pp. 7:1129-1159,1995.
- [8] M. En C. L. N. Oliva, "Sistemas de Procesamiento de Señales para el Análisis de Información Multidimensional". Tesis doctoral, Dep. Ing. Eléctrica, Electrónica de estado sólido, Cinevestav, 2007.
- [9] L.N.Oliva, O. A.C´A Rdenas, J.A. Moreno, L.M. Flores, F.Gomez., "Implementación de un Algoritmo ICA con circuitos CMOS Analógicos para el problema BSS". XII International Workshop. 2006.
- [10] Cardoso, J-F., "Infomax and maximum likelihood for blind source separation". *IEEE Signal processing Letters*. pp (4): 112-114 1997.
- [11] M. Paya, "Aislamiento térmico y acústico. International": Ediciones CEAC, 2004, pp. 17-23.

# Caracterización óptica de fibras ópticas sublongitud de onda

N.E. González-Sierra<sup>a\*</sup>, L.C. Gómez-Pavón<sup>a</sup>, R.G. Coxca-Gutiérrez<sup>a</sup>, G. F. Pérez-Sánchez<sup>b</sup>, P. Zaca-Morán<sup>b</sup>,  
A. Luis-Ramos<sup>a</sup>, J. M. Muñoz-Pacheco<sup>a</sup>.

<sup>a</sup>Facultad de Ciencias de la Electrónica, BUAP, Puebla, Pue., México.

<sup>b</sup>Instituto de Ciencias, BUAP, Puebla, Pue., México.

\*nancy.gonzalez@correo.buap.mx

**Resumen**—Las fibras ópticas sublongitud de onda exhiben propiedades interesantes tales como: intensos campos evanescentes, confinamiento óptico hermético y la posibilidad de miniaturizar sistemas; lo anterior sugiere la exploración de estas potencialidades para el desarrollo de dispositivos de sensado basados en luz. En este trabajo se presenta la respuesta óptica de microfibras ópticas fabricadas mediante la técnica *flame-brushing*. El experimento consiste en hacer propagar una onda coherente a una longitud de onda de 1550 nm a través de fibras ópticas con diámetros del orden de micrómetros. Los resultados obtenidos muestran que la transmisión en microfibras ópticas está relacionada con las dimensiones de las mismas. No se observó una respuesta en reflexión, lo que significa que gran parte de la señal es convertida en ondas evanescentes. Las fibras sublongitud de onda pueden ser utilizadas como dispositivos de sensado mediante campo evanescente.

**Keywords**—sublongitud de onda, campo evanescente, caracterización.

## I. INTRODUCCIÓN

DESDE hace 50 años, el sensado mediante fibra óptica ha sido una de las aplicaciones más poderosas y exitosas que conjuntan la fibra óptica y la tecnología de sensado. Las propiedades de las fibras ópticas, como lo es la inmunidad a la interferencia electromagnética, la posibilidad de transmitir luz a grandes distancias, y sus pequeñas dimensiones, las hacen ideales para el sensado de parámetros físicos, químicos y biológicos. Recientemente con el rápido progreso de la micro/nano tecnología se ha incrementado enormemente la demanda por sensores ópticos con mayor rendimiento y versatilidad, por lo que la miniaturización espacial ha sido una de las tendencias de los sensores de fibra óptica, puesto que reducir el tamaño de la estructura de sensado permite conferirle al sensor una respuesta más rápida, alta sensibilidad, bajo consumo de potencia y mayor resolución espacial, por esto las micro-nano fibras ópticas son componentes adecuados para tal propósito [1]–[4]. Históricamente, el primer reporte de fibras ópticas con diámetros de micrómetros fue realizado en el siglo XIX, por el físico británico Charles Vernon Boys, él logró obtener delgadas fibras de vidrio a partir de materiales fundidos mediante una técnica conocida como *flying arrow* [4], [5]. En 1959 Narinder S. Kapany, realizó una de las primeras aplicaciones para el guiado óptico, él logró transmitir imágenes a través de un arreglo de fibras ópticas micro y submicro métricas [6]. No fue

hasta la década de los 60' que Kao y Hockman proponen que la alta pureza de la fibra óptica traía consigo menores pérdidas [7], con lo que se consigue el desarrollo de las tecnologías mediante fibra óptica. Por varias décadas, las fibras ópticas con diámetros más grandes que la longitud de onda de luz transmitida, rápidamente encontraron un sinnúmero de aplicaciones incluyendo las comunicaciones ópticas, el sensado y la óptica no lineal. Sin embargo, en el 2003, L.Tong y otros coautores [8] demuestran experimentalmente las bajas pérdidas ópticas en fibras con diámetros mucho menores que la longitud de onda de la luz guiada, y se proponen las micro-nano fibras ópticas como los elementos básicos para el desarrollo de dispositivos fotónicos a pequeña escala. Cuando el diámetro de la fibra óptica es mucho menor que la longitud de onda de la luz que se propaga a través de la fibra óptica, se habla entonces de fibras ópticas sublongitud de onda, conocidas también en la literatura como alambres de fibra óptica sublongitud de onda, micro-nano fibras ópticas (nanofibras cuando el diámetro es menor a 1 micrómetro), fibras ópticas cónicas (*taper* en inglés) [5], entre otros. Las fibras ópticas sublongitud de onda son fabricadas mediante diversas técnicas, entre los diversos métodos resaltan los que forman parte del enfoque *top-down*, el cual consiste en primer lugar en el sometimiento del material a altas temperaturas para que sea posible manipularlo, y después la aplicación de una fuerza axial para reducir el material hasta dimensiones mucho menores que las originales. La técnica *flame-brushing* forma parte de las técnicas más reportadas en la literatura acerca de la fabricación de micro-nano fibras ópticas, dicha técnica forma parte del enfoque *top-down* para reducir un volumen relativamente grande a un volumen pequeño, esto es mediante el diseño cónico del material [1]. La técnica *flame-brushing* consiste en adelgazar la fibra óptica estándar; por lo que se conservan las dimensiones originales en los extremos de la fibra óptica, así es posible hacer conexiones directas a otros dispositivos ópticos. Para el guiado óptico, una excelente uniformidad geométrica así como una superficie lisa en las fibras ópticas sublongitud de onda son características críticas para garantizar que se tengan pocas pérdidas y una mejor relación señal a ruido, por ello la importancia en el proceso de fabricación de las fibras sublongitud de onda [3].



### I-A. Campos evanescentes en fibras ópticas sublongitud de onda

Las fibras ópticas sublongitud de onda en comparación con la fibra óptica convencional poseen un alto contraste entre el índice de refracción del núcleo y del medio circundante, además las dimensiones del núcleo son mucho menores que la longitud de la luz guiada, lo que hace que los alambres de fibra óptica sublongitud de onda presenten propiedades y características particulares cuando se transmite la luz en éstas, por ejemplo: alto confinamiento óptico, intensos campos evanescentes, alta no linealidad y bajas pérdidas cuando se interconecta con dispositivos fotoeléctricos [1]. El límite de difracción de la luz permite establecer las dimensiones mínimas que debe tener una fibra óptica para poder conducir la luz a través de ésta [9]. Sin embargo, las fibras ópticas sublongitud de onda pueden confinar la luz al límite de la difracción de la luz para longitudes que son solo determinadas por las pérdidas y comúnmente son varios ordenes de magnitud mayores que la longitud típica de Rayleigh [10] (éste régimen aplica diámetros menores que la longitud de onda de la luz guiada,  $2r \leq 0.2\lambda$  [11], donde  $r$  es el radio de la guía). Cuando una fibra óptica tiene un diámetro mucho menor que la longitud de onda de la luz guiada a través de la fibra óptica, gran parte de la potencia se propaga como un campo evanescente. Los campos evanescentes presentes en las fibras ópticas sublongitud de onda favorecen enormemente el sensado óptico [12], [13], y el acoplamiento de campos evanescentes ente las fibras ópticas sublongitud y otras guías de onda. El guiado óptico es una de las propiedades más importantes para aplicaciones ópticas y fotónicas. Para una fibra óptica estándar con un diámetro mucho mayor que la longitud de onda de la luz que se propaga en ésta, la intensidad del campo evanescente decae casi a cero fuera del revestimiento de la fibra óptica. Por lo tanto, la fibra óptica es insensible a los cambios en el medio circundante, lo que limita la interacción de la luz propagada en la fibra óptica y su entorno. Por otro lado, las fibras sublongitud de onda resultan ser más sensitivas en la superficie de la fibra óptica, debido a que la cantidad de ondas evanescentes y la intensidad del campo generado en la frontera de la fibra óptica es mucho mayor, lo que sugiere aprovechar dichas características para el diseño de plataformas de sensado miniaturizadas, más rápidas y sensitivas.

## II. DESARROLLO EXPERIMENTAL

Para el diseño y desarrollo de dispositivos fotónicos, conocer el comportamiento de la luz a través de las guías de onda es de suma importancia, ya que se determina las características físicas y geométricas que el dispositivo debe cumplir para determinadas aplicaciones, por esto, el desarrollo de este trabajo consistió en caracterizar ópticamente las fibras ópticas sublongitud de onda fabricadas mediante la técnica *flame-brushing*. A continuación se describe el desarrollo experimental:

### A. Fabricación de microfibras ópticas

Las microfibras ópticas se obtuvieron a partir de un sistema mecatrónico que se basa en la técnica *flame-brushing* [14].

Primero, a la fibra óptica se le retira el revestimiento en la sección que será sometida al calor, se limpia perfectamente con alcohol isopropílico y se coloca en las bases de sujeción. El proceso consiste en que la fibra óptica es calentada por el barrido de la flama de un mechero comercial de butano, una vez que el material del que está hecha la fibra óptica entra en un estado viscoso, las bases a la que se sujeta la fibra óptica ejercen una fuerza axial que permite la elongación de la misma. Una vez fabricadas las microfibras ópticas, éstas se almacenaron para su posterior caracterización óptica [15], y para evitar la contaminación en la superficie de la fibra óptica adelgazada.

### B. Caracterización de microfibras ópticas mediante la propagación de radiación láser

Para la caracterización óptica de microfibras ópticas, el experimento consistió en propagar una onda coherente a una longitud de onda de  $1550 \text{ nm}$  a través de las microfibras ópticas fabricadas, la fibra óptica bajo prueba se colocó en un portaobjetos de vidrio. En un extremo de la microfibra óptica se fusionó/empalmó el diodo láser y en el otro extremo de la fibra se detectó la potencia transmitida por medio de un sensor, los datos de la potencia de salida se visualizaron en el medidor de potencia. En la Figura 1 se muestra el arreglo experimental para la caracterización de las microfibras ópticas fabricadas, esto es mediante la propagación de radiación láser a través de ellas. Como se puede apreciar en la Figura 1, se detectó la potencia de salida de la microfibra óptica, de forma similar se analizó también la señal reflejada en el arreglo.

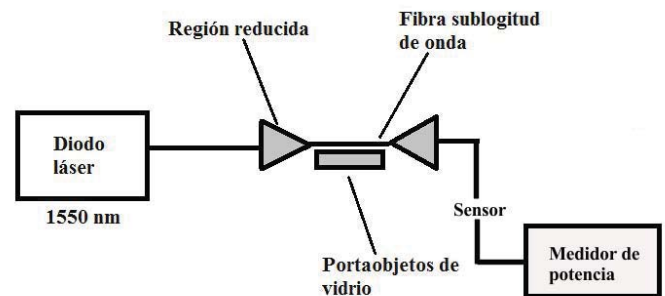


Figura 1. Esquemático para analizar la potencia de salida en fibras ópticas sublongitud de onda.

## III. RESULTADOS

Los resultados presentados en la Figura 2 muestran la variación de la potencia de salida en microfibras ópticas con respecto a la corriente suministrada por el diodo láser que emite a una longitud de onda de  $1550 \text{ nm}$ . La potencia transmitida en microfibras con diámetros de  $7, 8.6, 11, 17$  y  $25.3 \mu\text{m}$  es dependiente de las dimensiones de la microfibra óptica, se esperaba una relación proporcional al diámetro de la microfibra óptica, pero se observa que la transmisión no varía solo en función de ese parámetro, y se presume que tal

comportamiento es debido a la longitud de estiramiento de la microfibras óptica, lo que produciría un cambio en la potencia transmitida debido a la geometría de la región cónica, es decir, la región donde comienza a reducirse la fibra óptica de sus dimensiones originales hasta antes de la zona más delgada, en este caso, las microfibras de 7 y 11  $\mu\text{m}$  fueron fabricadas de manera que la longitud de estiramiento era mayor, a diferencia de las demás. Por lo que se resalta que la transmisión es aproximadamente igual en microfibras con diámetros de 8.6 y 17  $\mu\text{m}$ .

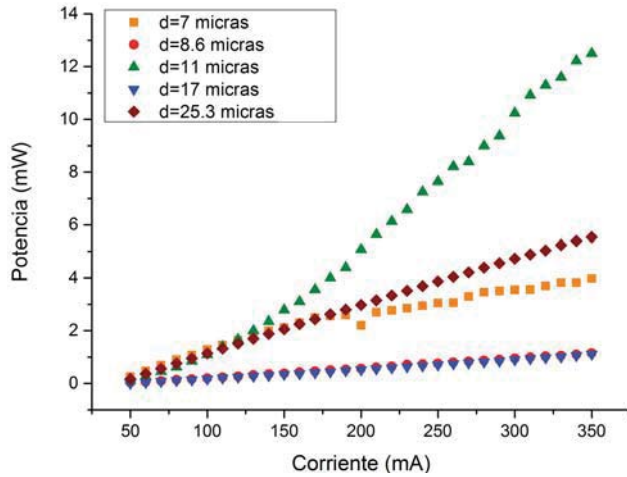


Figura 2. Resultados experimentales que indican la potencia a la salida de microfibras con diámetros de: 7, 8.6, 11, 17, 25.3  $\mu\text{m}$

En la Figura 3 se calculó el porcentaje de transmisión mediante la relación de la potencia de entrada y salida (Figura 2). Se aprecia que la transmitancia en microfibras ópticas con diámetros de 8.6, 17 y 25.3  $\mu\text{m}$  es aproximadamente constante con valores porcentuales de 10 %, 2 % y 2 % respectivamente, lo que indica el cambio proporcional entre la entrada y la salida, por otro lado, la transmitancia microfibras ópticas con diámetros de 7 y 11  $\mu\text{m}$  no es constante y se observa que la transmitancia tiene pendiente negativa y positiva respectivamente, debido a que la potencia de salida tiene cambios abruptos que promueven tal comportamiento.

Los resultados presentados en la Figuras 2 y 3 demuestran que existe una dependencia entre las características físicas de las microfibras ópticas y la potencia transmitida a través de éstas, en específico, en la Figura 3 el porcentaje en transmisión permanece constante en cierto valor porcentual para microfibras ópticas que presentaban una longitud de estiramiento menor a 1 cm (con diámetros de 8.6, 17 y 25.3  $\mu\text{m}$ ) comparado con las microfibras que tenían una longitud de estiramiento mayor a 1 cm (con diámetros de 7 y 11  $\mu\text{m}$ ). Con base a los resultados experimentales presentados, las microfibras ópticas fabricadas con diferente longitud de estiramiento serían propuestas según la aplicación, por ejemplo, si se requiriera que el dispositivo fotónico presentará una baja potencia y una transmisión constante, o se demandará la máxima transmisión (pérdidas mínimas).

Según lo reportado en la literatura, la potencia transmitida

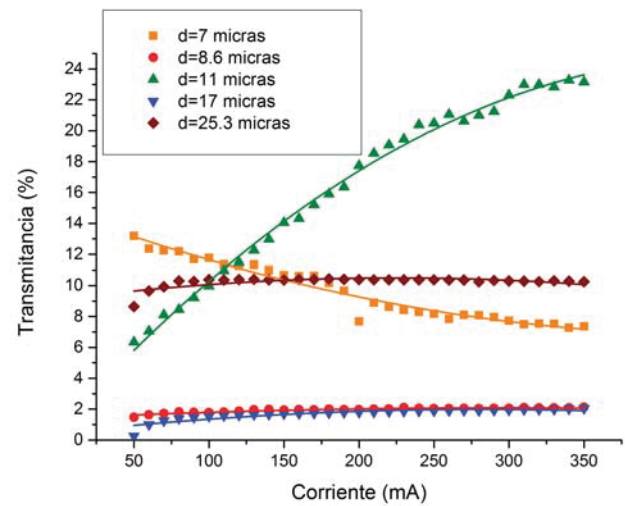


Figura 3. Resultados experimentales que indican la transmisión a la salida de microfibras con diámetros de: 7, 8.6, 11, 17, 25.3  $\mu\text{m}$

en microfibras ópticas esta principalmente relacionada con la geometría de éstas, en otras palabras, con el diámetro de la región más delgada, las regiones de transición y la longitud de estiramiento. En lo que respecta a las regiones de transición hacia la parte más delgada de la fibra óptica, éstas podrían ser de dos tipos: regiones de transición abruptas y regiones de transición con un decaimiento suave; en las primeras la pérdidas en transmisión serían mayores a diferencia de las segundas, esto es por el acoplamiento de los modos de propagación en la fibra óptica [16]–[18]. En este trabajo, se considera que un parámetro relevante en la transmitancia de las microfibras ópticas es la longitud de estiramiento, y se presume que la potencia transmitida es mayor debido a que las microfibras ópticas presentan una región de adelgazamiento más larga, lo que indicaría que las regiones de transición tienen un decaimiento suave. Por otro lado, se presume que las microfibras ópticas que presentaban una región de adelgazamiento menor, con una menor cantidad de potencia transmitida, estarían propagando la mayor parte de la potencia como un campo evanescente, en este caso, estos dispositivos responderían a cambios en su entorno de una manera más rápida, ideales como sensores altamente sensitivos por medio de campo evanescente.

Cabe mencionar que además de las pérdidas en transmisión asociadas a las características físicas de las microfibras ópticas, el que la superficie de las fibras ópticas esté contaminada o no presente uniformidad y paredes lisas aumenta las pérdidas en transmisión [5]. En este trabajo, las microfibras ópticas fueron analizadas en el microscopio óptico y el microscopio electrónico de barrido (SEM), este equipo nos permitió estimar el diámetro de las microfibras, así como su morfología, y se observó que las microfibras ópticas fabricadas presentaban paredes lisas.

#### IV. CONCLUSIÓN

Es importante predecir las condiciones físicas que un dispositivo fotónico debe cubrir para determinada aplicación, por lo

que resulta relevante considerar la propagación de la radiación óptica en guías de onda a escala nanométricas, en este caso, en microfibras ópticas, debido a que el continuo avance de la nanotecnología demanda dispositivos que consuman menos potencia, sean más pequeños y tengan una respuesta más rápida, por lo que las microfibras ópticas son candidatas para el desarrollo de sensores fotónicos. Los resultados demuestran que la fabricación de las microfibras ópticas así como su caracterización mediante la propagación de radiación láser a través de ellas es importante, debido a que se pueden establecer los parámetros físicos que hacen que el dispositivo sea propuesto como un sensor vía campo evanescente.

#### AGRADECIMIENTOS

Esta investigación fue parcialmente soportada por el proyecto VIEP-BUAP-2016.

#### REFERENCIAS

- [1] J. Lou, Y. Wang y L. Tong, "Microfiber optical sensors: A review", *Sensors*, Vol. 14, pp. 5823-5844, 2014.
- [2] X. Guo, Y. Ying y L. Tong, "Photonic nanowires: From subwavelength waveguides to optical sensors", *Accounts of Chemical Research*, Vol. 47, No. 2, pp. 656-666, 2014.
- [3] X. Wu y L. Tong, "Optical microfibers and nanofibers", *Nanophotonics*, Vol. 2(5-6), pp. 407-428, 2013.
- [4] L. Tong y M. Sumetsky, "Subwavelength and nanometer fabrication, properties and applications", *IEEE*, Vol. 12, pp. 2728-2730, 2012.
- [5] L. Tong, F. Zei, X. Guo y L. Tong, "Optical microfibers and nanofibers: a tutorial", *Elsevier*, Vol. 285, pp. 4641-4647, 2012.
- [6] N. S. Kapany, "High-resolution fibre optics using sub-micron multiple fibres", *Nature*, Vol. 184, pp. 881-883, 1959.
- [7] K. C. Kao, G. A. Hockham, "Dielectric-fibre surface waveguide for optical frequencies", *Proc. IEE*, Vol. 113, pp. 1151-1158, 1966.
- [8] L. Tong, R. R. Gatass, J. B. Asbcom, S. He, J. Lou, M. I. Maxwell y E. Mazur, "Subwavelength-diameter silica wires for low-loss optical wave guiding", *Nature*, Vol. 426, pp. 816-819, 2013.
- [9] Fundación Eiticotaku (Dic. 2009), Nanoantenas ópticas: más allá del límite de difracción, Magazine, [En línea], (Enero 2016).
- [10] G. Y. Chen, M. Ding, T. P. Newson y G. Brambilla, "A review of microfiber and nanofiber based optical sensors", *The Open Optics Journal*, Vol. 7, pp. 32-57, 2013.
- [11] A. Ashkin, J. M. Dziedzic, J. E. Bjorkholm y S. Chu, "Observation of a single-beam gradient force optical trap for dielectric particles", *Opt. Lett.*, Vol. 11, pp. 288-290, 1986.
- [12] L. Zhang, F. X. Gu, J. Y. Lou y L. M. Tong, "Fast detection of humidity with a subwavelength-diameter fiber taper coated with gelatin film", *Opt. Express*, Vol. 16, pp. 13349-13402, 2008.
- [13] A. Coillet, B. Cluzel, G. Vienne, P. Grelu y F. Fornel, "Near-field characterization of glass microfibers on a low-index substrate", *App. Phys B*, Vol. 101, pp. 291-296, 2010.
- [14] J. A. Rodríguez Sánchez, L.C Gómez Pavón, A. Luis Ramos, G. F. Pérez Sánchez, J. M Muñoz Pacheco, O.G. Félix Beltrán y P. Zaca Morán, "Sistema mecatrónico para la fabricación de nanoalambres de fibra óptica", *SOMI*, pp. 1-7, 2014.
- [15] R.G Coxca Gutiérrez, "Fabricación de micro-nano fibras ópticas para el desarrollo de dispositivos fotónicos", Tesis de maestría, Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, Puebla, México, 2016.
- [16] J. D. Love, W. M. Henry, W. J. Stewart, R. J. Black, S. Lacroix, y F. Gonthier, "Tapered single-mode fibres and devices. Part I: Adiabaticity criteria", *IEE Proc.*, Vol. 138, pp. 343-354, 1991.
- [17] T. A. Birks y Y. W. Li, "The shape of fiber tapers", *J. Lightwave Technol.*, Vol. 10, pp. 432-438, 1992.
- [18] J. Villatoro, D. Monzón Hernández, y E. Mejía, "Fabrication and modeling of uniform-waist single-mode tapered optical fiber sensors", *Appl. Optics*, Vol.42, No. 13, pp. 2278-2283, 2003.

# Nodo sensor infrarrojo para un arquitectura multicapa

V.H.Garcia, R.O. Gonzalez, J.E. Martinez

Dpto. de Ing. en Sistemas Computacionales  
Escuela Superior de Cómputo del IPN  
Av. Juan de Dios Batiz s/n, Col. Lindavista, 07738,  
Ciudad de México  
[vgarciao@ipn.mx](mailto:vgarciao@ipn.mx), [rortegag@ipn.mx](mailto:rortegag@ipn.mx),  
[jmartinez0809@alumno.ipn.mx](mailto:jmartinez0809@alumno.ipn.mx)

J.C. Sosa

Centro de Investigación en Ciencia Aplicada y Tecnología  
del IPN  
Cerro Blanco 141, Col Colinas del Cimatario, 76090,  
Querétaro, Qro.  
[jcsosa@ipn.mx](mailto:jcsosa@ipn.mx)

**Resumen**—Este artículo describe el diseño de un nodo sensor, para una arquitectura de dos capas en una red de sensores inalámbrica. Esta red de sensores es aplicada a domótica para el monitoreo en casa-habitación. La primer capa contempla la comunicación entre el nodo base y nodos sensores que utilizan controladores digitales de señales como unidad de procesamiento. La segunda capa contempla la comunicación entre los nodos sensores con los controladores digitales de señales y nodos sensores denominados infrarrojos, por su tipo de comunicación. Los nodos sensores infrarrojos utilizan un microcontrolador PIC12LF1840 de ocho terminales de bajo consumo como unidad de procesamiento. Para la comunicación entre los nodos de la segunda capa se utiliza un diodo y un módulo receptor infrarrojo, los cuales usan una portadora de 56KHz con una trama de 18 bits para el envío de hasta 16 códigos de función diferentes. Con estos códigos se identifica el estado del nodo infrarrojo y se solicita la información de los sensores conectados a él. Los nodos sensores infrarrojos usan un sensor digital de efecto hall para monitorear la apertura o cierre de ventanas y puertas, y un sensor analógico para el monitoreo de temperatura.

**Palabras clave**—Red de sensores inalámbrica, nodo sensor, comunicación infrarroja, domótica, microcontrolador.

## I. INTRODUCCION

Una red de sensores inalámbricos (WSN - Wireless Sensor Network) es una red inalámbrica que consiste en nodos de sensores autónomos implementados en zonas de interés que tienen en común características como: procesamiento de datos, capacidad de almacenamiento, interfaces de comunicación inalámbrica y consumo de energía limitada. Estas redes son utilizadas para monitorear y controlar diversos tipos de aplicaciones en diferentes tipos de ambientes [1].

Una WSN consiste de un número de nodos de sensores (algunas decenas, cientos o incluso miles) trabajando al mismo tiempo para monitorear una región y obtener datos de los fenómenos físicos que sean de interés para el administrador. Hay dos tipos de WSN's: estructuradas y no estructuradas.

Una WSN no estructurada es aquella que contiene una densa colección de nodos de sensores colocados en el lugar de interés de una manera aleatoria donde cada nodo puede o no ser implementado de una manera "Ad Hoc". Una vez colocados, se dejan para que presenten y reporten sus funciones. Su mantenimiento, así como la administración

de su conexión y detección de fallas es difícil ya que raramente se conoce su ubicación [2].

En una WSN estructurada, todos o la mayoría de los nodos son colocados en posiciones planeadas. La ventaja de las redes estructuradas es que con pocos nodos se puede cubrir un área mayor debido a que son colocados en posiciones estratégicas con un mínimo de supervisión y con bajo costo en su mantenimiento [3].

Cuando se tiene una WSN estructurada, las posiciones de cada nodo sensor son conocidas y la información que obtiene cada nodo sensor es enviada a un nodo base para ser procesada y/o también puede ser enviada a una puerta de enlace para que la información sea colocada en internet o en una intranet. Por lo tanto, los dispositivos que conforman una red inalámbrica de sensores son [4]:

**Un nodo base (sinknode).** Es un dispositivo que puede ser concebido como una laptop que recibe datos de una red de sensores o de un dispositivo aún más pequeño como un microcontrolador. Puede proporcionar la funcionalidad de una puerta de enlace quien crea un puente entre la red de sensores y una red externa como por ejemplo internet [5, 6].

**Una puerta de enlace inalámbrica (wirelessgateway).** Traduce contenidos y protocolos de los miembros de una red para que los demás miembros puedan entender los mensajes enviados. Una puerta de enlace inalámbrica provee una amplia gama de servicios como el establecimiento de sesiones entre clientes y servidores, seguridad, compresión/descompresión de información y manejo de fallas [7].

**Nodo sensor inalámbrico.** Es un dispositivo autónomo que es utilizado para medir un fenómeno físico mediante sus sensores, procesar la información necesaria y es capaz de transmitir los resultados de manera inalámbrica. La figura 1 muestra cada módulo que constituye a un nodo sensor inalámbrico.

Actualmente las aplicaciones de las redes de sensores inalámbricas son muy amplias. Como ejemplos, tenemos las siguientes [8]:

- Sistema de Aguas [9]
- Seguridad y emergencia [10]
- Control industrial [11]
- Agricultura [12]
- Ganadería [13]

- Domótica y edificios inteligentes [14]
- Salud [15]

La domótica es un conjunto de diferentes tecnologías aplicadas al monitoreo, control y automatización de sistemas y dispositivos en la vivienda. Los principales objetivos de la domótica son mejorar la seguridad personal y patrimonial de la vivienda, aumentar el confort y tener una gestión eficiente del uso de la energía.

Un sistema domótico está formado por una WSN. Esta red, es capaz de recolectar información, procesarla y emitir órdenes a unos actuadores o salidas. La arquitectura de una aplicación domótica puede tener dos tipos de comunicación: la comunicación interior y la comunicación exterior, tal como se muestra en la Fig. 1.

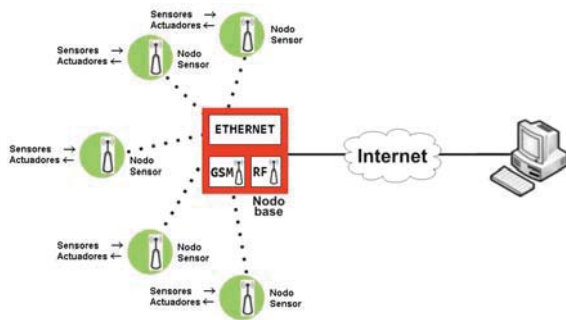


Fig. 1. Arquitectura básica de un sistema domótico.

En la arquitectura mostrada en la Fig. 1 se muestra un Nodo base que tiene la función de puerta de enlace para poder monitorear y/o controlar de forma remota los sensores colocados en una casa habitación. También tiene la función de comunicarse de forma inalámbrica con los nodos sensores mediante una red de área personal inalámbrica (WPAN – Wireless Personal Area Network) bajo el estándar IEEE 802.15.4. Para formar una WPAN se requiere el uso de un transceptor que opere con el estándar, presentando algunas desventajas al momento de la implementación de la red:

- El transceptor suele incorporar un cierto número de terminales dependiendo de la interfaz de comunicación con la que dispone, normalmente usan la interfaz de periféricos serial (SPI – Serial Peripheral Interface), esto conlleva a usar microcontroladores de al menos 14 terminales para interconectarlo.
- El consumo de corriente en modo de bajo consumo del transceptor es mayor en comparación con otros medios de transmisión como el infrarrojo.
- Si solo se requiere monitorear unos o dos sensores digitales o analógicos, conectados al nodo sensor, el número de terminales necesarias para los sensores son pocas. Lo que aumenta el número de terminales a usar en el microcontrolador del nodo sensor, en realidad esta determinado por el transceptor.

En este artículo se propone el desarrollo de un nodo sensor con un módulo de comunicación infrarrojo para reducir el

consumo de corriente y el número de terminales en el microcontrolador usado como nodo de procesamiento. Este nodo se incorpora en una WSN de arquitectura multicapa.

## II. ARQUITECTURA DEL SISTEMA

La arquitectura general del sistema se muestra en la Fig. 2 y está basada en [16-21].

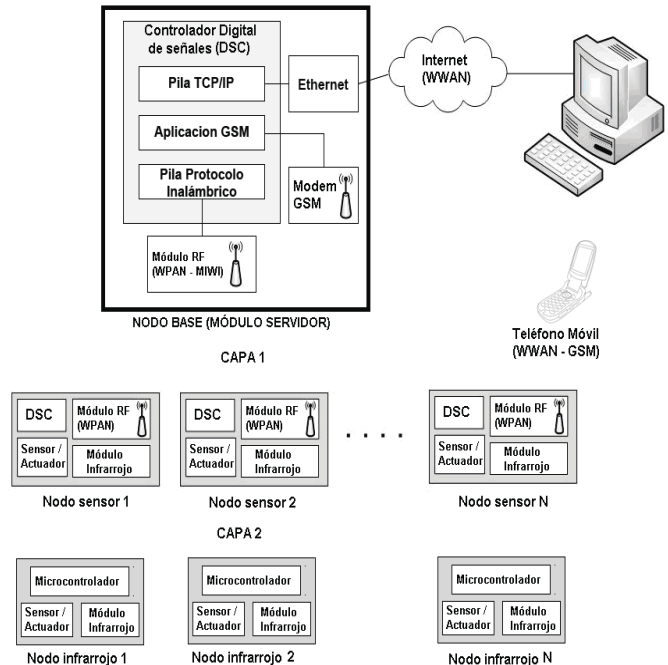


Fig. 2. Arquitectura general del sistema.

La arquitectura de la WSN aplicada a domótica incorpora los siguientes elementos:

### A. Nodo Base

Este nodo está basado en el controlador digital de señales DSPIC33EP512GP502 de la compañía Microchip, como unidad de procesamiento. Este nodo contiene implementada la pila de protocolo TCP/IP para la comunicación a través de Internet usando el controlador Ethernet ENC624J600 y la pila de protocolo MIWI [27] para la comunicación a través de WPAN usando el transceptor MRF24J40MB [16].

El Sistema Global para Comunicaciones Móviles (GSM – Global System for Mobile Communications) se ha usado en diversas aplicaciones de monitoreo y control [22-26] usando microcontroladores de 8 bits. En la arquitectura propuesta, el nodo base utiliza un modem GSM que permite realizar la comunicación entre la red de sensores y el usuario usando un teléfono móvil. El modem GSM se configura, sin usar una computadora personal, mediante comandos AT, los cuales son enviados desde el DSC para su inicialización, configuración, envío y recepción de mensajes. El modem GSM utiliza una interfaz de Transmisión y Recepción Asíncrona Universal (UART – Universal Asynchronous Receiver Transmitter) de comunicación para el envío y recepción de los comandos AT desde el DSC. Con este módulo de comunicación el usuario

puede controlar y monitorear los sensores usando el servicio de mensajes cortos (SMS – Short Message Service).

**B. Nodo Sensor**

Este nodo también está basado en el controlador digital de señales DSPIC33EP512GP502 de la compañía Microchip, como unidad de procesamiento. Este nodo contiene implementada la pila de protocolo MIWI para la comunicación a través de la WPAN con el nodo base. Este nodo puede manejar sensores digitales, sensores analógicos y sensores digitales con una interfaz de comunicación como UART, el bus de Interconexión de Circuitos Integrados (IIC – Inter Integrated Circuit Bus) y SPI.

El nodo sensor contiene un módulo de comunicación infrarrojo con el cual recibimos y enviamos información a los nodos infrarrojos. Este módulo infrarrojo utiliza un diodo y un módulo receptor infrarrojo, los cuales usan una portadora de 56KHz con una trama de 18 bits para el envío de hasta 16 códigos de función diferentes. Con estos códigos se identifica el estado del nodo infrarrojo y se solicita la información de los sensores conectados a los nodos infrarrojos.

**C. Nodo infrarrojo**

Este nodo esta basado en el microcontrolador PIC12LF1840 de la compañía Microchip, como unidad de procesamiento. Este nodo, también contiene un módulo de comunicación infrarrojo para la comunicación con los nodos sensores.

**III. NODO INFRARROJO**

Las partes que componen el Nodo infrarrojo son:

- La unidad de procesamiento.
- El módulo de comunicación infrarrojo.
- Sensores a monitorear

**A. Unidad de Procesamiento**

El microcontrolador PIC12LF1840 tiene las siguientes características [28]:

- Arquitectura RISC de 8 bits
- Oscilador externo de hasta 32 Mhz e interno de hasta 16MHz.
- Convertidor Analógico Digital (ADC – Analog to Digital Converter) de 10 bits con 4 canales.
- Módulos: SPI, IIC y EUSART
- Encapsulado de ocho terminales con 6 disponibles de entrada/salida.

Las terminales del microcontrolador tienen funciones compartidas entre los diferentes periféricos, por lo que se pueden realizar algunas combinaciones de funciones dependiendo de los sensores o actuadores a conectar. Las seis terminales están asignadas a las funciones mostradas en la tabla 1.

TABLA I. FUNCIONES DE TERMINALES EN EL PIC12LF1840

Terminal	Función
1	VDD (1.8V – 3.6V)
2	RX – Receptor
3	TX – Transmisor
4	MCLR – Master Clear
5	Pin Configurable 2 – PC2
6	Pin Configurable 1 – PC1
7	Pin Configurable 0 – PC0
8	VSS – Tierra

Las terminales 1 y 8 son alimentación, en las terminales 2 y 3 se encuentra la señal de recepción y transmisión del UART del microcontrolador, por lo que podemos conectar el módulo infrarrojo, una interfaz USB a UART a través de un FT232 [30] o un módulo con el estándar Bluetooth [29]. Las terminales 5, 6 y 7 son pines configurables que pueden funcionar como pin de entrada/salida digital, entrada analógica, módulo UART, IIC y SPI. Las terminales se muestran en la Fig. 3.

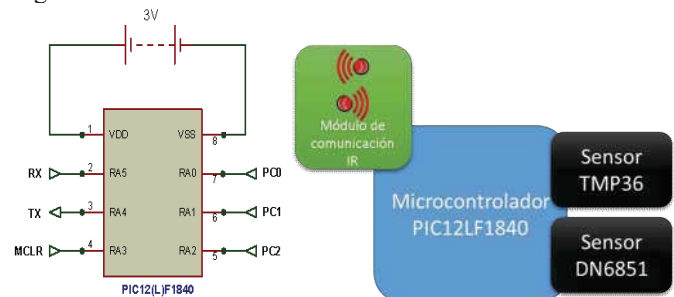


Fig. 3. Distribución de terminales y sensores en el PIC12LF1840.

Las diferentes funciones de los pines configurables se muestran en la tabla II.

TABLA II. FUNCIONES DE LOS PINES CONFIGURABLES EN EL PIC12LF1840

Funciones	PC2	PC1	PC0
Digitales – D	D	D	D
Analógicos - A	A	A	A
Combinación D/A	D/A	D/A	D/A
UART	D	RX	TX
UART	A	RX	TX
IIC	SDA	SCL	D
IIC	SDA	SCL	A
SPI	MISO	SCK	MOSI

Para este trabajo se uso un módulo infrarrojo, para el envío y recepción de datos en la capa 2 con los nodos sensores, usando las terminales RX y TX. Un sensor de efecto Hall y un sensor de temperatura.

El programa principal del microcontrolador que realiza el envío y recepción de datos usando el módulo infrarrojo y el monitoreo de los sensores de efecto hall y temperatura se muestra en la Fig. 3.

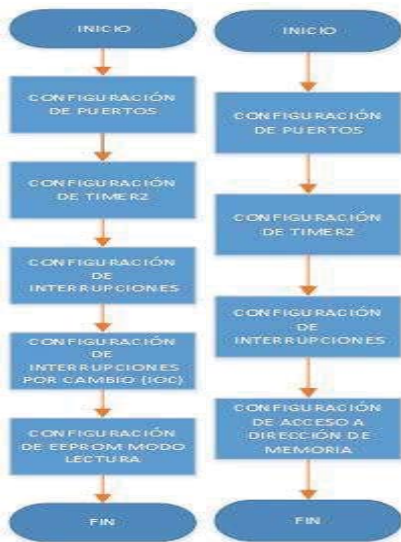


Fig. 4. Programa principal del nodo infrarrojo y base.

El TIMER 2 del microcontrolador se usa para generar la señal portadora del módulo infrarrojo a través del sistema de interrupciones. Se utiliza una interrupción por cambio para la recepción de los datos en este módulo.

Se utiliza el TIMER 1 para generar un reloj de tiempo real y la memoria EEPROM del microcontrolador se usa para guardar el estado de los sensores.

**B. Módulo de comunicación infrarrojo**

Este módulo esta formado por un diodo emisor infrarrojo con una configuración para alcanzar 4 metros de distancia, tal como se muestra en la Fig 5. El envío de la trama consiste en un tren de pulsos cuadrados de 56 KHz. la denominada portadora. Un nivel "0" ese envía como 500µs de silencio, el "1" como 500µs de pulsos.

Para el módulo receptor infrarrojo se utiliza el modelo TSOP34156 [33]. Tiene un fotodetector y preamplificador en el mismo encapsulado, además de un filtro interno. Es un receptor para sistemas de control remoto por infrarrojos. La señal de salida es demodulada directamente y puede ser decodificada por un microcontrolador.

Para la parte de comunicación se diseño una trama con un tamaño fijo de 18 bits. Esta trama esta formada por dos bits de cabecera para identificar que la trama es la que se espera, cuatro bits para código de función lo que permite tener 16 funciones y finalmente 12 bits de datos, esto se puede apreciar en la Fig. 5.

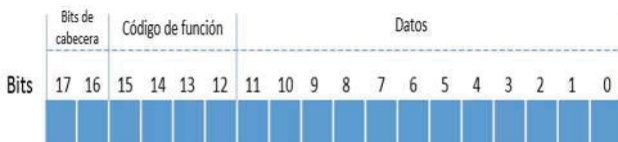


Fig. 5. Trama de comunicación para el nodo infrarrojo.

Con los 4 bits de código de función se pueden definir hasta 16 funciones diferentes. Para este trabajo se definieron solo 6 funciones. Dependiendo de la función, cada bit del campo de datos en la trama tiene un significado específico. Las tramas definidas se muestran en la tabla III.

TABLA III. TRAMAS DEFINIDAS PARA LA COMUNICACIÓN

BITS	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1	0
Trama de Tiempo	1	1	0	0	0	1	BC*	H*	H	H	H	H	M*	M	M	M	M	M
Trama de Identificación	1	1	0	0	1	0	IN*	IN	IN	IN	IN	IN	IN	IN	UN*	UN	UN	UN
Trama de Descripción	1	1	0	0	1	1	X	X	X	X	X	X	X	X	T*	T	T	T
Trama de Lectura	1	1	0	1	0	0	X	X	IS*	IS	D	D	D	D	D	D	D	D
Trama de no Información	1	1	0	1	0	1	X	X	X	X	X	X	X	X	X	X	X	X
Trama de Petición	1	1	1	1	1	1	X	X	X	X	IN	IN	IN	IN	IN	IN	IN	IN

Trama de tiempo. Esta trama es usada para saber el tiempo en el que la lectura fue registrada. Son necesarios 11 bits para el envío de estos datos; 5 bits para las horas y 6 para los minutos, el tiempo registrado se enviará en el formato de 24 horas. El bit 11 es el bit de control de tramas, es utilizado para conocer el número de tramas a recibir, toma el valor de '1' si se recibirán 8 tramas o '0' si se recibirán 4. Ver Fig. 6.

Trama de identificación. Esta trama es usada para enviar la dirección del nodo dentro de la red y su ubicación dentro del recinto, para determinar el origen de los datos recibidos. Se utilizan 12 bits para el envío de los datos; 8 bits para la dirección del sensor, ya que pueden ser agregados más nodos a la red siendo un total de 256 nodos y 4 para la ubicación del nodo.

Trama de descripción. Esta trama es usada para darle a conocer al nodo sensor los diversos tipos de sensores que operan en el nodo esclavo, se utilizan 4 bits.

Trama de lectura. Esta trama es usada para el envío de los datos obtenidos por el sensor. Se utilizan 10 bits para el envío de los datos, 8 bits para la lectura del sensor y 2 bits para identificar el tipo de sensor del cual son los datos. Ver Fig. 6.

Tama de no información. Esta trama es usada cuando el nodo esclavo no tiene información que enviar al nodo coordinador, esto sucede cuando no se ha detectado un cambio en los sensores de nodo.

Trama de petición. Esta trama es usada para el envío de la petición de datos obtenidos por el sensor, la petición se envía del nodo sensor hacia los nodos infrarrojos. Se utilizan 12 bits para el envío de los datos, 4 bits con valores no importa y 8 bits con la dirección del nodo, siendo un total de 256 direcciones posibles.

**C. Sensores**

En el nodo infrarrojo se utiliza un sensor de efecto Hall digital modelo DN6851 [31] conectado a PC2, por lo que éste se configura como digital. Este sensor se usa con la finalidad de detectar la apertura y cierre de puertas y/o ventanas.

También se configura un sensor de temperatura analógico modelo TMP36 [32] conectado a PC1, por lo que éste se configura como analógico. Este sensor es acondicionado al rango dinámico del ADC y es encargado de monitorear cambios de temperatura en el ambiente. Su voltaje de salida esta dado por:

$$V_{out} = \left( 10 \frac{mV}{^{\circ}C} * temp^{\circ}C \right) + 500mV \quad (1)$$

Los sensores se muestran en la Fig. 3.

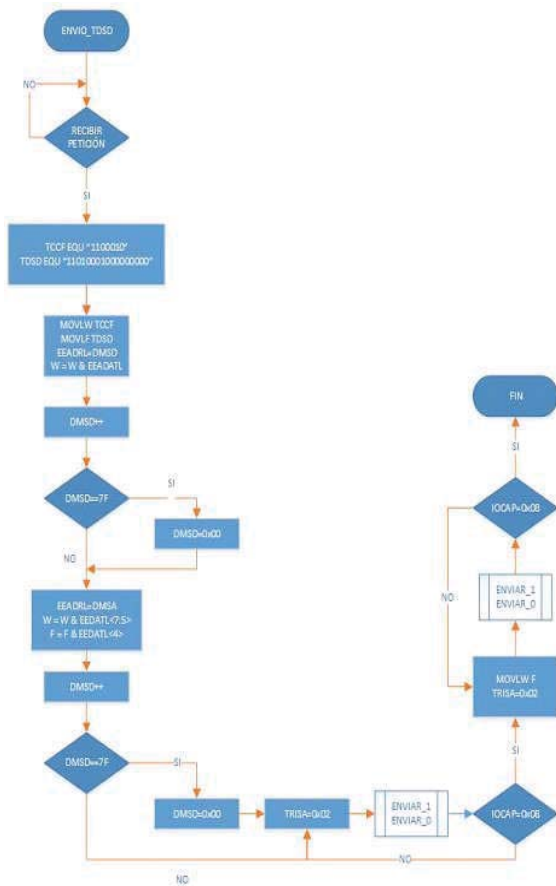


Fig. 6. Envío de trama de tiempo y lectura digital.

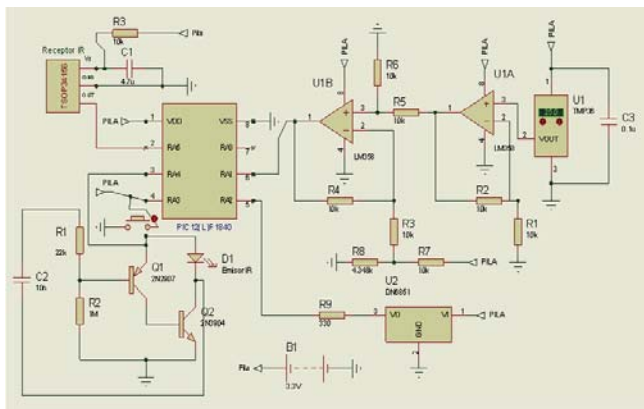


Fig. 7. Diagrama esquemático completo del nodo infrarrojo.

#### IV. PRUEBAS Y RESULTADOS

El nodo infrarrojo fue probado de manera independiente a la arquitectura multicapa usando un nodo base con un DSPIC30F3013 como unidad de procesamiento, el módulo infrarrojo para la comunicación y un módulo GSM para el

monitoreo remoto. El módulo GSM utilizado en esta aplicación es el GSM Click, el cual es una tarjeta de expansión que hace uso del factor de forma mikroBUS™ y contiene un módulo GSM/GPRS modelo GL865-QUAD de la compañía Telit y un socket para tarjeta SIM. Éste módulo se comunica con el DSPIC30F3013 a través de UART y puede ser utilizado ya sea a 3.3 o 5 volts. La interfaz UART de este dispositivo está configurada por defecto a una velocidad de 9600 baudios y una trama con un bit de inicio, ocho bits por dato y un bit de paro [21].

El módulo GL865-QUAD ofrece la pila del protocolo 3GPP y soporta las bandas de 850/900/1800/1900 MHz.

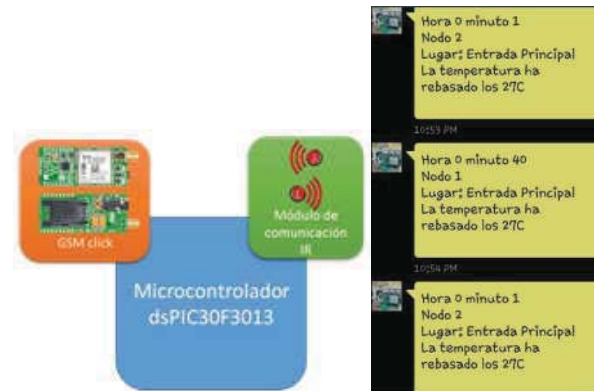


Fig. 8. Diagrama a bloques del nodo base de prueba y de los resultados.

El nodo infrarrojo se diseño modularmente en 3 tarjetas de circuitos impresos (PCB – Printed Circuit Board) con la finalidad de realizar fácilmente la adaptación de nuevos sensores y módulos de comunicación. Los módulos son:

Un PCB contiene el PIC12LF1840 con un conector donde se coloca el módulo de comunicación y los sensores. Se tiene un segundo conector que se utiliza para programar el microcontrolador usando la interfaz ICSP [34].

Otro PCB contiene el módulo de comunicación infrarrojo y el último PCB contiene los sensores de temperatura y de efecto hall. Esto se muestra en la Fig. 9.



Fig. 9. Nodo infrarrojo y nodo base.

El nodo base le envía peticiones al nodo infrarrojo mediante las tramas para la comunicación y lectura de los sensores. Los datos obtenidos de los sensores son enviados al usuario usando el modem GSM cuando se genera un evento. En el caso del sensor de temperatura cuando rebasa un umbral



de 27°C se manda un mensaje de aviso al móvil del usuario. En el caso del sensor de efecto hall se manda un mensaje cuando se detecta la apertura de una puerta o ventana. Esto se muestra en la Fig. 8. El programa principal del nodo base se muestra en la Fig 4.

## V. CONCLUSIONES

El nodo sensor infrarrojo permite realizar una buena comunicación en un espacio máximo de 4 metros sin obstáculos.

Al utilizar el módulo infrarrojo de comunicación solo se usan dos terminales del PIC12LF1840, permitiendo usar un microcontrolador de ocho terminales de bajo consumo y costo.

La comunicación se realiza mediante una trama de 18 bits la cual permite definir hasta 16 códigos de función y hasta 256 nodos infrarrojos mediante la trama de petición.

Este nodo se puede incorporar en una segunda capa de comunicación en una WSN para enviar datos de sensores cuya distancia es corta evitando así transeptores más sofisticados y costosos.

El nodo infrarrojo al ser modular permite la adaptación de otros tipos de sensores digitales, analógicos o con una interfaz de comunicación y módulos de comunicación como una interfaz USB a UART a través de un FT232 o un módulo con el estándar Bluetooth.

## AGRADECIMIENTOS

Se agradece al IPN por el apoyo brindado a la realización de este trabajo a través del financiamiento del proyecto SIP 20151421.

## REFERENCIAS

- [1] J Fernandez-Berni and R Carmona Galán. Vision-enabled WSN Nodes: State of the Art. ELSEVIER. 2012.
- [2] B. Xu, S. Hischke, and B. Walke, "The Role of Ad hoc Networking in Future Wireless Communications," IEEE conference publications, vol. 2, pp. 1353 - 1358, 2003.
- [3] Biswanath Mukherjee, Dipak Ghosal Jennifer Yick, "Wireless sensor network survey," ELSEVIER, no. 52, pp. 2292-2330, Abril 2008.
- [4] Yang Lv and Yu Tian, "Design and application of sink node for Wireless Sensor Network," IEEE 2nd International Conference on Industrial and Information Systems, pp. 487-490, 2010.
- [5] Lili Yang, "Determining Sink Node Locations in Wireless Sensor Networks," IEEE International Conference on Systems, Man, and Cybernetics, pp. 3400-3404, Octubre 2006.
- [6] Shahabuddin Muhammad, Zeeshan Furqan, and Ratan Guha, "Wireless Sensor Network Security: A Secure Sink Node Architecture," IEEE International Performance, Computing, and Communications Conference, pp. 371-376, Abril 2005.
- [7] Amjad Umar, "Mobile Computing and Wireless Communications," in Mobile Computing and Wireless Communications, nge solutions, Ed.: NGE Solutions, 2004, ch. 1.7.1.
- [8] Libelium. Libelium Comunicaciones Distribuidas S.L. [Online]. [http://www.libelium.com/es/top\\_50\\_iot\\_sensor\\_applications\\_ranking](http://www.libelium.com/es/top_50_iot_sensor_applications_ranking)
- [9] Seena Verma and Prachi, "Wireless Sensor Network application for water quality monitoring in India," IEEE National Conference on Computing and Communication Systems, 2012.
- [10] Kewey Sha and Weisong Shi, "Using WSN for Fire Rescue applications: Requirements and Challenges," IEEE International Conference on Electro/information Technology, pp. 239 - 244, 2006.
- [11] Gao Baolu, Xiong Shibo, and Xu Zhanwei, "The Application of Wireless Sensor Networks in Machinery Fault Diagnosis," IEEE Computer Society, pp. 315 - 318, 2010.
- [12] Takashi Togami and Kyosuke Yamamoto, "A Wireless Sensor Network in a vineyard for smart viticultural management," SICE Annual Conference, pp. 2450 -2454, Septiembre 2011.
- [13] Kae Hstang Kwong and Konstantinos Sasloglou, "Adaptation of Wireless Sensor Network for farming industries," IEEE, 2009.
- [14] Adrian Valentin Nedelcu, Florin Sandu, Mihai Machedon-Pisu, Marian Alexandru, and Petre Ogrutan, "Wireless-based Remote Monitoring and Control of Intelligent Buildings," IEEE, p. 6, 2009.
- [15] Stanislaw Rajba and Teresa Rajba, "Wireless Sensor Networks in Application to Patients Health Monitoring," IEEE Symposium on Computational Intelligence in Healthcare and e-health (CICARE), pp. 94 - 98, 2013.
- [16] V. H. García, et. al. "Módulo servidor para domótica usando TCP/IP. 1er Congreso Internacional de Robótica y Computación (CIRC)". La Paz, Baja California Sur, México. 2013.
- [17] V. H. García, et. al. "Red inalámbrica de comunicación para el monitoreo y control en una casa habitación". VII Congreso Internacional de Telemática y Telecomunicaciones (CITTEL). La Habana, Cuba. 2012.
- [18] V. H. García, et. al. "Sistema para domótica inalámbrico". XXIV Congreso Nacional y X Congreso Internacional de Informática y Computación ANIEI 2011 (CNCIIC-ANIEI). Colima, Colima. 2011
- [19] V. H. García, et. al. "Desarrollo de un sistema SCADA para casa habitación". 32 Congreso Internacional de Ingeniería Electrónica. (ELECTRO). Chihuahua, Chihuahua. 2010.
- [20] V. H. García, et. al. "Desarrollo de un sistema SCADA inalámbrico para casas habitación". 5th International Congress on Electronics and Biomedical Engineering, Computation and Informatics.(CONCIBE SCIENCE). Guadalajara. 2009.
- [21] V. H. García, et. al. "Prototipo de sistema para control y monitoreo de una red de sensores mediante comunicación GSM aplicada a domótica". Pistas educativas, No 108, pp 1718-1739, Octubre 2014.
- [22] B. Ramamurthy. "Development of a Low-Cost GSM SMS-Based Humidity Remote Monitoring and Control system for Industrial Applications", International Journal of Advanced Computer Science and Applications. Vol. 1, No. 4, October 2010.
- [23] S.R. Khan, "Design and Implementation of Low Cost Home Security System using GSM Network". International Journal of Scientific & Engineering Research. Volume 3, Issue 3, March -2012.
- [24] V. Madan, "GSM-Bluetooth based Remote Monitoring and control system with automatic light controller". International Journal of Computer applications. Volume 46-No 1, May -2012.
- [25] A. Alheraish, "Design and implementation of home Automation system". IEEE Transactions on Consumer Electronics, Volume 50, No 4, November 2004.
- [26] A.W. Ahmad, "Implementation of Zigbee-GSM based Home Security Monitoring and Remote Control System". Circuits and Systems (MWSCAS), 2011 IEEE 54th International Midwest Symposium on, Aug 2011.
- [27] Microchip Technology Inc. *MiWi(TM) Development Environment Help*. Chandler, Arizona. 2016. <http://www.microchip.com>.
- [28] Microchip Technology Inc. *PIC12(L)F1840 Data Sheet*. Chandler, Arizona. 2016. <http://www.microchip.com>.
- [29] Bluetooth SIG. *Bluetooth Core Specification*. Kirkland. 2016. <http://www.bluetooth.com>.
- [30] FTDI Chip. *FT232R USB UART IC Datasheet*. 2016. <http://www.ftdichip.com>
- [31] Panasonic, Inc. *DN6851 Data Sheet*. 2016. <http://www.semicon.panasonic.co.jp/>.
- [32] Analog Devices, Inc. *TMP36 Data Sheet*. 2016. [www.analog.com](http://www.analog.com).
- [33] Vishay Semiconductors, Inc. *TSOP34156 Data Sheet*. 2016. [www.vishay.com](http://www.vishay.com).
- [34] Microchip Technology Inc. *ICSP*. Chandler, Arizona. 2016. <http://www.microchip.com>

# Caracterización de micro-nano fibras ópticas fabricadas mediante un sistema mecatrónico desarrollado a partir de la técnica flame brushing.

R.G. Coxca-Gutiérrez<sup>a</sup>, L.C. Gómez-Pavón<sup>a</sup>, N.E. González-Sierra<sup>a\*</sup>, G. F. Pérez-Sánchez<sup>b</sup>, P. Zaca-Morán<sup>b</sup>,  
A. Luis-Ramos<sup>a</sup>, J. M. Muñoz-Pacheco<sup>a</sup>.

<sup>a</sup>Facultad de Ciencias de la Electrónica, BUAP, Puebla, Pue., México.

<sup>b</sup>Instituto de Ciencias, BUAP, Puebla, Pue., México.

\*rafael.coxca@alumno.buap.mx

**Resumen**—Los alambres sublongitud son de gran interés debido a las propiedades ópticas que presentan, tales como grandes campos evanescentes, alta no linealidad, extrema flexibilidad y configurabilidad además de poseer bajas pérdidas de interconexión. En este trabajo se presenta la caracterización de micro-nano fibras ópticas por medio de la técnica flame-brushing, que ha sido desarrollada mediante un sistema mecatrónico con el que se han obtenido alambres de hasta 800nm. La caracterización de las fibras ópticas se lleva a cabo a partir de parámetros como la velocidad de estiramiento, longitud de calentamiento y precalentamiento los cuales son modificados a través de un software además de establecer las dependencia en el proceso de fabricación. Los resultados obtenidos establecen los parámetros óptimos para la fabricación de micro-nano fibras ópticas los cuales pueden ser utilizados en aplicaciones de sensado, interferometría y otros.

**Keywords**—Fibra óptica, micro-nano fibras, alambres sublongitud, flame brushin, sistema mecatrónico.

## I. INTRODUCCIÓN

EN los últimos años se han presentado un sinnúmero de ideas para fabricar nuevos tipos de fibras, las cuales tendrían una gran cantidad de aplicaciones en varios o en casi todas las áreas de desarrollo científico y tecnológico. Una de estas ideas es de reducir el diámetro de la fibra óptica a partir de un volumen material, para confinar la longitud de onda de la luz en el núcleo de tal manera de sobrepasar el límite de la difracción [1], [2]. Cuando el diámetro del núcleo de la fibra óptica desciende por debajo de la longitud de onda de la señal luminosa transmitida, se dice que la fibra es un alambre sublongitud de onda (o micro-nano fibra óptica).

Históricamente desde el siglo XIX era posible realizar estiramiento de fibras finas altamente uniforme [3]. Sin embargo el uso de estas fibras como guías de onda óptica no se aplicó durante un largo tiempo. No fue hasta la década de 1960, cuando la teoría de guía de onda óptica había sido bien establecida, se comenzó a investigar las propiedades de las guías de onda ópticas de fibras previamente estiradas por lo que comenzaron a jugar un papel más importante [4].

En el año 2003 se demostró experimentalmente que las micro nano fibras ópticas tenían bajas pérdidas, ya que en los procesos de fabricación presentan dimensiones originales de

fibra óptica en los extremos de entrada y salida, debido a su forma cónica [5].

Las micro-nano fibras ópticas han atraído cada vez más atención debido a la amplia gama de componentes que se han implementando en función de sus propiedades y configurabilidad que facilitan dispositivos altamente compactos y de geometría compleja. Entre los cuales se han explotado la creación resonadores de alto factor Q, guías de átomos, generación de supercontinuos, acopladores ópticos, filtros y amplificación de campo evanescente que favorecen a la creación de dispositivos de sensado [5].

En su mayoría la fabricación de alambres sublongitud se basa en técnicas que tienen un enfoque de arriba hacia abajo (top-down) como consecuencia de reducir materiales de gran volumen a fibras finas por medio de la forma de su dibujo cónico como se muestra en la Figura 1. Estas técnicas se pueden aplicar a una variedad de materiales como vidrio y plástico los cuales ofrecen cierta viscosidad adecuada para el dibujo cónico [6].

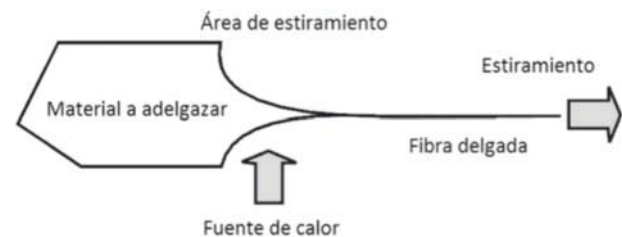


Figura 1. Diagrama para el adelgazamiento de fibra óptica.

En particular una de las técnicas que nos permite obtener micro-nano fibras es la técnica flame-brushing el cual es el sistema típico de estiramiento por medio del barrido de una fuente de calor en la sección media de una fibra estándar para después aplicar en cada uno de los extremos una fuerza de arrastre que provoca el adelgazamiento de la sección media de la fibra en la Figura 2 se muestra el diagrama de la técnica. El proceso continúa con una fuerza de estiramiento constante aplicada en los dos extremos, y la fibra se estira gradualmente hasta que se alcanza la longitud deseada, generalmente la

fuentes de calor esta basada en un soplete de gas de hidrogeno debido a su limpieza, y fácil control, que proporcionan la posibilidad de tener la temperatura suficientemente alta para el estiramiento de la fibra [7].

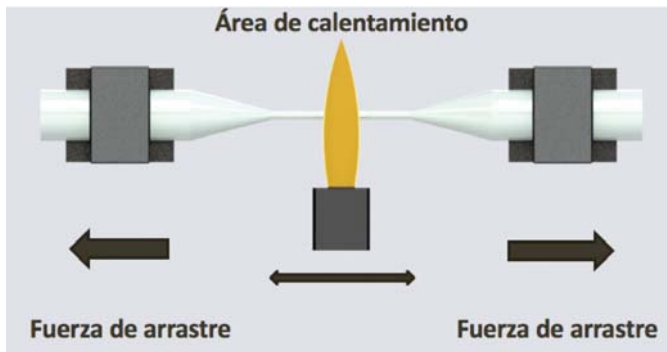


Figura 2. Esquemático del estiramiento mediante la técnica flame brushing.

En este trabajo nos proponemos a determinar los parámetros óptimos para la fabricación de micro-nano fibras ópticas mediante el sistema mecatrónico diseñado y construido por el Grupo de Sistemas Fotónicos y Nanoóptica de la facultad de ciencias de la electrónica de la BUAP. El sistema se basa en la técnica flame-brushing y se proponen definir: El valor o rango de la velocidad de estiramiento, la longitud de calentamiento y el precalentamiento de la fibra óptica con el fin de garantizar la repetibilidad de la manufactura de micro-nano fibras con diferentes diámetros.

## II. DESCRIPCIÓN DEL SISTEMA

El sistema mecatrónico consta de una estructura mecánica y una interfaz electrónica que controla los motores que proporcionan la fuerza de arrastre y ubicación de la fuente de calor por medio de un software de control. En las siguientes secciones se describen cada una de estas etapas:

### II-A. Estructura mecánica.

Está compuesta por una base la cual barre una fuente de calor a lo largo de la fibra, el barrido de la base se realiza por medio de un motor, además la fuente de calor esta sujeta a otro motor con el que se puede cambiar la ubicación de calentamiento de la fuente de calor. El sistema de estiramiento lo componen dos bases, estas bases se encuentran sobre una estructura de soporte las cuales contienen una plataforma de desplazamiento conectado a un tornillo que proporcionan un desplazamiento micrométrico, en la Figura 3 se muestra el diagrama a bloques del sistema mecatrónico desarrollado por el grupo de sistemas Fotónicos y Nanoópticas.

### II-B. Estructura mecánica.

Tomando en cuenta las características de los motores que desplazan el sistema de estiramiento, la base de barrido de la flama y el cambio de ubicación de la flama, se implementaron

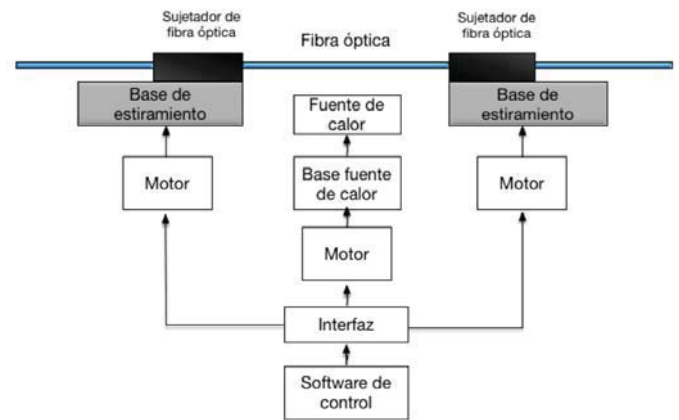


Figura 3. Diagrama a bloques del sistema mecatrónico desarrollado a partir de la técnica flame brushing.

controladores con las capacidades de cambiar el sentido, frecuencia y activación del motor, los controladores establecen una comunicación por medio del puerto USB, para llevar a cabo dicha comunicación USB, se incluyó un microcontrolador con capacidades suficientes de funcionar como interfaz y transmitir las señales definidas por el usuario. La interfaz de hardware consiste de un microcontrolador PIC18F4550 programado para funcionar con el protocolo de comunicación USB. El PIC seleccionado realiza la comunicación entre la interfaz electrónica desarrollada y la computadora, así como el envío y recepción de los datos del software de control desarrollado en LabVIEW.

En la Figura 4 se presenta el diagrama de flujo general de la interfaz electrónica desarrollada, la cual deberá realizar el cambio de la velocidad de estiramiento, frecuencia de barrido de la fuente de calor y cambio de ángulo de la fuente de calor.

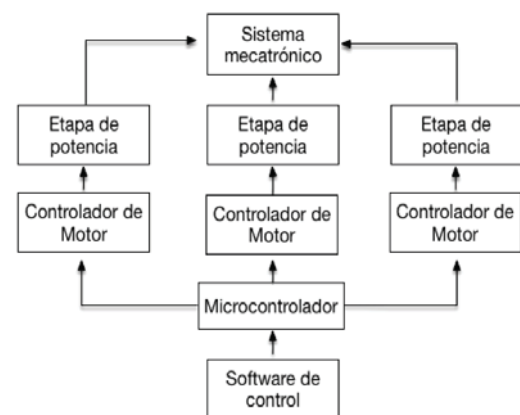


Figura 4. Diagrama de flujo de la interfaz electrónico general de la interfaz electrónica.

### II-C. Software de control.

El software se diseño a través de la plataforma Labview la cual es utilizada para la automatización de procesos, por

medio de dicho software el usuario tiene opción de modificar las siguientes secciones:

- **Parámetros de estiramiento;** Contiene tres diferentes formas de activar los motores, la opción “Motores síncronos” con lo cual los motores desplazan las bases a la misma velocidad y distancia, la opción “Motores no sincronizados” tiene la opción de tener diferentes velocidades y distancias desplazadas, por ultimo la opción “Rutina de tiempo” con la que se pueden realizar rutinas en con hasta diez cambios de velocidades y diferentes intervalos de tiempo.
- **Parámetro de precalentamiento;** Con esta sección se tiene la posibilidad de realizar un barrido de la fuente de calor antes de iniciar el proceso de estiramiento además también se puede controlar la frecuencia de barrido.
- **Ángulo de la flama;** En esta sección contiene la opción de poder realizar un cambio de la posición, dirección, distancia y tiempo en el que la fuente de la fuente de calor cambia.

Además en el software de control contiene un paro de emergencia y un sistema de monitoreo con el cual se puede observar la fabricación de las micro-nano fibras ópticas.

En Figura 5 se presenta el panel frontal del software de usuario del sistema, donde se encuentran integradas cada una de estas secciones.

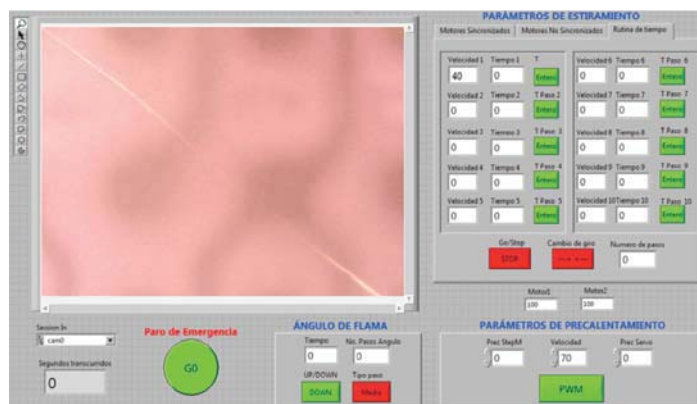


Figura 5. Software de control utilizado para la técnica flame brushing.

### III. FABRICACIÓN DE MICRO-NANO FIBRAS ÓPTICAS

Antes de iniciar con las pruebas experimentales se prepara la fibra óptica, el procedimiento para preparar la fibra óptica consiste en remover el recubrimiento a lo largo de la longitud de la fibra óptica a utilizar. Para remover el recubrimiento se utiliza un micro pelador de fibra óptica, quedando la fibra con revestimiento y el núcleo.

Una vez realizado el proceso de preparación de la fibra óptica se procede mediante el software de control a colocar las bases de desplazamiento lo más cerca posible con el fin de aprovechar la máxima distancia a estirar, enseguida se coloca la fibra en los sujetadores de fibra óptica.

Parámetros controlables	
Velocidad de Estiramiento	5 $\mu\text{m/s}$ -195 $\mu\text{m/s}$
Precalentamiento	20s-80s
Frecuencia de barrido	70Hz-180Hz
Longitud de calentamiento	1.2-3.4cm

Cuadro I. INTERVALOS DE LOS PARÁMETROS VARIADOS.

Se ingresan los valores al software de control en las secciones de parámetros de estiramiento, parámetros de calentamiento y a ángulo de la flama, en el Cuadro I se muestra los intervalos de los parámetros. Estos parámetros los mantenemos fijos a lo largo de la prueba en donde se observa el comportamiento que tiene la fibra a largo de su fabricación, al término de la prueba se anota la descripción de los diferentes cambios que la fibra tuvo durante la prueba estos datos son: desplazamiento de las bases, comienzo de deformidad de la fibra, tiempo total hasta que la fibra se fractura.

Se repiten las pruebas esperando observar los mismos cambios y así poder establecer los mejores parámetros en donde la fibra fabricada presente los mejores resultados. A las fibras fabricadas con las mejores rutinas se les realiza una caracterización microscópica en la cual se obtiene su diámetro y longitud, se realizara la repetibilidad de la rutina con la que espera obtener los mismos resultados.

### IV. RESULTADOS

Como se mencionó anteriormente el software de control contiene la sección “Parámetros de precalentamiento”, En las pruebas realizadas se observó que este parámetro depende de la velocidad de las bases de desplazamiento ya que a mayor tiempo de precalentamiento la fibra se vuelve cada vez mas viscosa lo que ocasiona una pronta deformación por la fuente de calor en el caso de que de que las bases se desplacen lento, de esta manera si se deja un largo tiempo de precalentamiento es necesario velocidades iguales de estiramiento ya que se puede ocasionar una deformación de la fibra durante la fabricación, o por el contrario, los tiempos de precalentamiento cortos están asociados a velocidades pequeñas al inicio de la rutina, lo que implica un lento y gradual adelgazamiento de la fibra. Se observó que los tiempos de precalentamiento en un intervalo de 25 a 90 segundos son apropiados para la fabricación con velocidades constantes de hasta 50  $\mu\text{m/s}$ , además en la misma sección se hicieron variaciones del barrido de la fuente de calor entre 90 y 120Hz, se observa un adelgazamiento menor cuando este parámetro aumenta.

Suponiendo que al aumentar la longitud de desplazamiento del sistema de estiramiento, se reduce el diámetro de la fibra óptica. Se selecciona la opción rutinas de tiempo de la sección parámetros de estiramiento en la se van variando la velocidad a lo largo del proceso de fabricación con el fin de hacer una reducción gradual hasta que la longitud y el diámetro de la fibra sea el deseado, el cambio de velocidad se realiza cuando la fibra comienza a deformarse por la presión de la flama, variaron en un intervalo de 7  $\mu\text{m/s}$  a 50  $\mu\text{m/s}$  y el tiempo de fabricación de la fibra ronda entre los 900s.

Como resultado de la rutina se obtuvo fibras con varias reducciones significativas las cuales fueron caracterizadas a

partir del microscopio electrónico de barrido(SEM). En la Figura 6 se observa una fibra sublongitud de onda que tiene una reducción de aproximadamente  $25\mu\text{m}$  y una longitud aproximada de  $0,8\text{mm}$ .

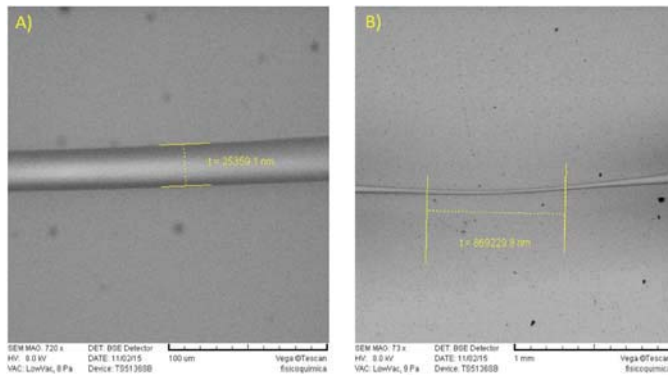


Figura 6. Fibra sublongitud de onda vista desde el microscopio óptico de barrido(SEM): parte menos adelgazada (lado izquierdo) y longitud de la parte mas delgada (lado derecho).

En la Figura 7 se muestra una segunda prueba donde se aumentó la longitud de las bases de desplazamiento que estiran a la fibra sublongitud, se observa que tiene dos regiones reducidas de aproximadamente  $16\mu\text{m}$ , presentaba una longitud aproximada de  $16\text{mm}$ .

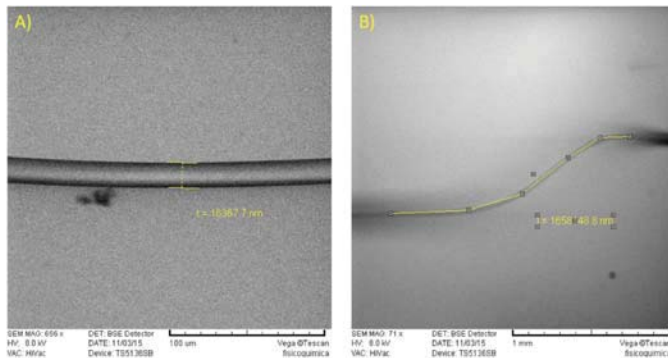


Figura 7. Fibra sublongitud de onda vista desde el microscopio óptico de barrido(SEM): parte menos adelgazada (lado izquierdo) y longitud de la parte mas delgada (lado derecho).

En la Figura 8 se observa una fibra sublongitud de onda fabricada a partir de una rutina de tiempo corto, en la que se comienza adelgazando muy rápido y se termina de estirar a una velocidad media, debido a que el comienzo de la rutina el adelgazamiento es rápido se establece un tiempo de precalentamiento de 90s.

## V. CONCLUSIONES.

Las fibras fabricadas mediante la técnica flame brushing se analizaron por medio del microscopio óptico y microscopio electrónico de barrido donde se pudo observar las características que presentaban cada vez que se modificaba algún

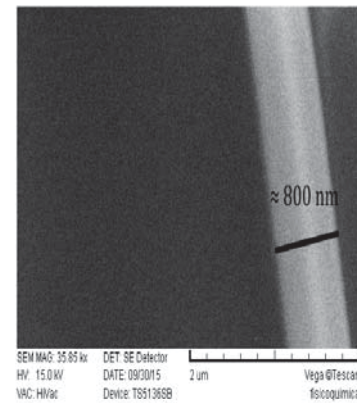


Figura 8. Parte más delgada de la fibra sublongitud de onda vista desde el microscopio óptico de barrido(SEM).

parámetro que intervienen en el proceso. Se observó que mediante una longitud de calentamiento corta se pueden obtener fibras con longitudes cortas y diámetros pequeños. Además, alargando el tiempo de precalentamiento en el proceso se pudo notar que la fibra puede soportar velocidades altas durante el proceso de fabricación. Considerando las pruebas anteriores, se tiene como objetivo construcción de rutinas con las cuales se pueda garantizar obtener alambres sublongitud con las características necesarias para diversas aplicaciones.

## AGRADECIMIENTOS

Esta investigación fue parcialmente soportada por el proyecto VIEP-BUAP-2016.

## REFERENCIAS

- [1] F. L. Kien, V. I. Balykin, K. Hakuta, "Angular momentum of light in an optical nanofiber", *Phys. Rev. A* 73, 053823 (2006)
- [2] S. Kawata M. Ohtsu. M. Irie, *Nano-Optics* (2002) Ed. Springer.
- [3] C. V. Boys, On the production, properties, some suggested uses of the finest threads, *Phil. Mag.* 23, 489-499 (1887)
- [4] Jesus Castellon Uribe. *Optical Fiber Sensors: An Overview*. INTECH Open Access Publisher (2012)
- [5] L. M. Tong, R. R. Gattass, J. B. Ashcom, S. L. He, J. Y. Lou, M. Y. Shen, I. Maxwell, E. Mazur, Subwavelength-diameter silica wires for low-loss optical wave guiding, *Nature* 426, 816-819 (2003)
- [6] L. Tong, M. Sumetsky, "Fabrication of MNFs", *Subwavelength and Nanometer Diameter Optical Fibers*, Ed. Springer (2010)
- [7] M. Sumetsky, "How thin can a microfiber be and still guide light?", *Opt. Lett.* Vol. 31, Issue 7, pp. 870-872 (2006)

# Modelo de datos para el análisis de información georreferenciada sobre biodiversidad aplicado en un entorno de aplicaciones móviles

Alberto González Espinoza  
Maestría en Sistemas Computacionales  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
Email: alberto.gonzalez@openmailbox.org

Luis Armando Cárdenas Florido  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
Email: armando.cardenas@gmail.com

Guillermo Martínez Flores  
Departamento de Oceanología  
Centro Interdisciplinario de Ciencias Marinas  
Instituto Politécnico Nacional  
La Paz, Baja California Sur, México  
Email: gmflores@ipn.mx

**Abstract**—Las bases de datos de biodiversidad, almacenan información histórica sobre la variedad de organismos que constituyen la vida sobre la tierra, recabada por la comunidad científica a lo largo de más de 250 años de investigación. Existen además, diferentes proyectos que integran bases de datos de acceso abierto sobre biodiversidad que pueden ser explotadas por medio de servicios web. Para poder analizar estos datos y obtener información significativa para el usuario, se desarrolló un modelo de datos multidimensional, utilizando el sistema SQL Server Analysis Services, y se aplicaron técnicas de bodegas de datos (data warehouses) para crear una herramienta para la visualización de información desde diferentes perspectivas o dimensiones (hipercubo). Este modelo permite analizar la información desde 6 dimensiones básicas: lugar, tiempo, grupo de organismos, región, colecciones y base del registro. Además, debido a que los usuarios de los sistemas sobre biodiversidad requieren acceso a la información para su trabajo de campo (comúnmente en sitios sin acceso a Internet), se desarrolló una aplicación móvil para dispositivos Android, capaz de consultar la información primaria sobre biodiversidad de los servicios web de acceso abierto, así como acceder al hiper cubo creado, para generar informes resumidos sobre la biodiversidad en las regiones geográficas de interés, almacenando estos datos en la memoria interna del dispositivo para consultarlos fuera de línea.

## I. INTRODUCCIÓN

Podemos entender el concepto de biodiversidad como la variedad de genes, especies y ecosistemas que constituyen la vida sobre la tierra [1]. De esta forma, dentro del campo de los sistemas de información, los datos sobre la biodiversidad se pueden organizar en tres niveles: genético/molecular, especies y ecosistemas [2]. La información de las especies, se almacena en bases de datos en forma de Registros Primarios de Biodiversidad (PBR, por sus siglas en inglés), cada uno de los cuales describe la aparición de un espécimen en un momento y punto geográfico determinados [3]. En la actualidad, el mayor esfuerzo para integrar una base de datos centralizada a nivel mundial, conformada por registros primarios de biodiversidad, ha sido el proyecto de la Infraestructura de Información Mundial sobre Biodiversidad (GBIF, por sus siglas en inglés), brindando acceso a esta información para su explotación a través de Internet [3].

Además, existen otros proyectos que brindan información complementaria a los registros primarios de biodiversidad,

como la Enciclopedia de la Vida (EOL por sus siglas en inglés), en la cual podemos encontrar una colección de páginas con información biológica validada de una gran cantidad de especies de seres vivos [4].

Toda esta información, conforma el conocimiento primario sobre la biodiversidad para la comunidad científica, pero además, al analizarse adecuadamente, puede ser aplicada para la resolución de problemáticas concretas, como la designación de áreas naturales protegidas, diseño de políticas para el manejo de la tierra y el combate de especies invasoras [5].

Las bodegas de datos o data warehouses, son de gran ayuda para trabajar con esta información histórica acumulada [6], ya que en éstas se pueden almacenar todos los datos masivos de apariciones de organismos, y aplicar un modelo de datos multidimensional que nos permita analizarla y comprenderla más fácilmente.

Los modelos de datos multidimensionales, permiten analizar la información de la bodega de datos desde diferentes perspectivas, visualizando los datos en un espacio  $n$ -dimensional, usualmente llamado un cubo de datos o hiper cubo [6], el cual permite sintetizar la información y organizarla según las perspectivas o dimensiones definidas, de forma que resulta mucho más comprensible para el usuario.

La georreferenciación es el proceso de relacionar la información con una ubicación geográfica, utilizando un sistema de ubicaciones geoespaciales (coordinadas angulares como longitud y latitud) [7]. Puesto que los datos de este proyecto se encuentran georreferenciados, es posible agregar dimensiones espaciales al modelo, pudiendo definir polígonos de coordenadas geográficas en la base de datos, que correspondan a regiones de interés para el usuario que analiza la información. De esta forma, se pueden extraer y clasificar a todos los ejemplares dentro de una región geográfica, y analizarlos en conjunto de manera independiente del resto de la base de datos.

Finalmente, las aplicaciones para dispositivos móviles, nos ofrecen nuevas oportunidades para desarrollar innovaciones, dentro del terreno de los sistemas de información geográfica sobre biodiversidad [8]. Gracias a su capacidad de almacenar información fuera de línea y a su GPS incorporado, podemos acceder a la información consultada previamente estando en

localidades fuera del alcance de Internet, así como ubicar los lugares de aparición de los ejemplares, utilizando el sistema de geoposicionamiento del dispositivo móvil.

## II. DATOS Y MÉTODOS

El desarrollo de este proyecto constituyó cinco etapas principales:

- 1) Extracción, transformación y carga de la información
- 2) Desarrollo del modelo de datos
- 3) Diseño de las consultas multidimensionales
- 4) Creación del cliente móvil
- 5) Conexión de la aplicación móvil con el hiper cubo

### A. Extracción, transformación y carga de la información

Como fuente de datos primaria para el desarrollo del modelo de datos multidimensional, se utilizó la base de datos de Registros Primarios de Biodiversidad de GBIF. De esta forma, se pueden utilizar datos reales históricos para probar el hiper cubo desarrollado.

De esta gran base de datos formada por más de 500 millones de registros, se extrajo solo un fragmento que cubriera geográficamente la región noroeste del territorio mexicano, utilizando la herramienta para exportación de datos que brinda el sitio web oficial del proyecto, se delimitó un polígono geográfico como se muestra en la fig. 1.

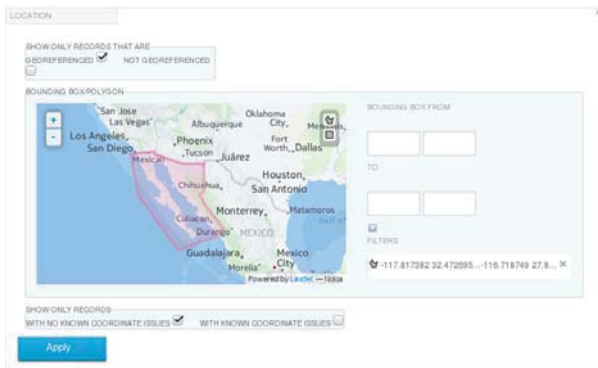


Fig. 1. Captura de pantalla de la herramienta de exportación de datos del sitio oficial de GBIF. Con esta herramienta se delimitan las regiones que serán exportadas.

Esta información se extrajo en un conjunto de archivos en formato CSV y XML, que fueron importados dentro de una base de datos intermedia en MS SQL Server. Después, fue necesario transformar estos datos para poder analizarlos al momento de aplicar el modelo multidimensional, de modo que se eliminaron datos no relevantes, se convirtieron los datos de las fechas, se corrigieron los nombres de lugares y se incorporaron nuevas tablas, para clasificar geográficamente los datos dentro de países y de regiones geográficas, delimitadas por polígonos de coordenadas geográficas, de forma que se complementara la información original con una nueva dimensión. El resultado en esta etapa se puede apreciar en el diagrama de la figura 2.

Para la etapa de carga, se utilizó el motor de análisis de datos MS SQL Server Analysis Services, a través del software MS Visual Studio 2010 Shell. La bodega de datos se creó a partir de un nuevo proyecto dentro de este entorno, y para la

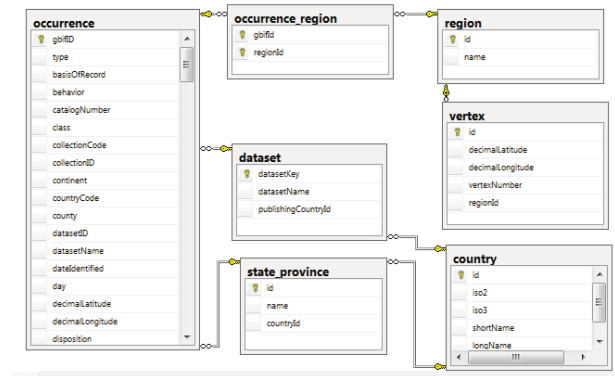


Fig. 2. Diagrama entidad-relación de la base de datos construida para la creación de la bodega de datos. En este se pueden apreciar la inclusión de los catálogos de colecciones, países y regiones geográficas.

carga de la información se utilizó una conexión al servidor de base de datos MS SQL Server por medio del SQL Native Client 11.1 (SQLNCLI11.1). Una vez conectada a la fuente de datos (DataSource) se creó una vista de la fuente de datos (Data Source View), en la que se seleccionaron las tablas que serían incorporadas a la bodega de datos. Estas fueron: occurrence, region, occurrence\_region, vertex, state\_province, country y dataset. El resultado de la carga se puede apreciar en la fig. 3.

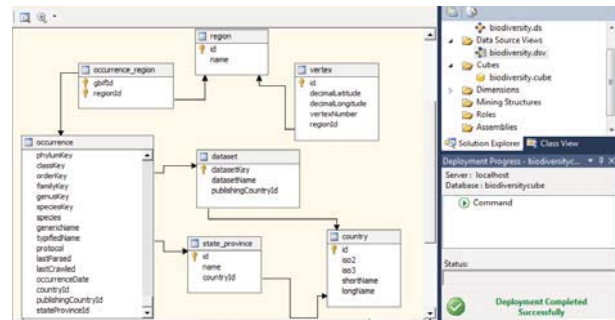


Fig. 3. Visualización de las tablas cargadas dentro de la bodega de datos creada con la herramienta Analysis Services Multidimensional and Data Mining.

### B. Diseño del modelo multidimensional

Un modelo multidimensional típico cuenta con al menos una tabla de hechos, que es la parte central del diseño, ya que representa el objeto de nuestro análisis. Estos hechos son clasificados dentro de diferentes dimensiones y poseen al menos una métrica de interés para el usuario [6]. Para este proyecto, el punto medular es la información acerca de las apariciones de organismos, por lo que la tabla de hechos se obtuvo a partir de la tabla occurrence.

La métrica fundamental para este diseño es el número de ejemplares registrados en los eventos de investigación. Esta métrica se calcula directamente a partir del número de registros contabilizados en la tabla occurrence. Por ejemplo, la métrica para la especie *Scarus ghobban*, sería igual al número de registros que hacen referencia a esta especie.

El tipo de esquema utilizado para representar la tabla de hechos y las dimensiones de este modelo, es el esquema en

estrella, ya que se cuenta con una tabla de hechos central de la cual dependerán todas las dimensiones diseñadas (figura 4).

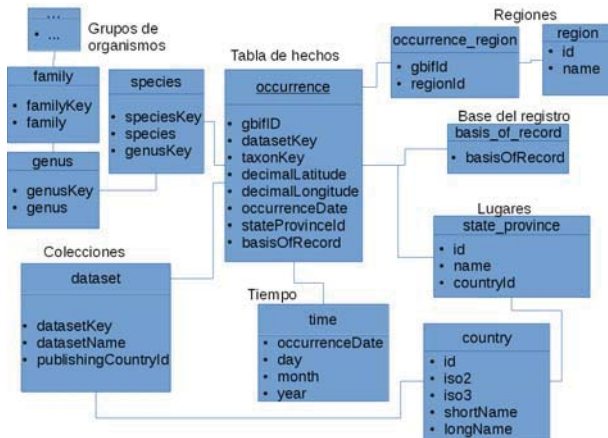


Fig. 4. Esquema del modelo multidimensional diseñado, en el que se aprecian las seis grandes dimensiones que describen la información: grupo de organismos, lugares, regiones, colecciones, base del registro y tiempo.

Para este modelo de datos, se crearon seis dimensiones para el análisis de la información:

- 1) Taxonomy. El grupo de organismos al que pertenece el ejemplar.
- 2) Place. Localidades geográficas a los que hacen referencia los registros.
- 3) Dataset. Colección biológica donde está registrado el ejemplar.
- 4) Region. Áreas geográficas formadas por un polígono de coordenadas dentro de las cuales puede haberse encontrado un ejemplar.
- 5) Time. Día, mes y año de la aparición del ejemplar.
- 6) BasisOfRecord. Indica la forma en la que se obtuvo el registro del organismo: ya sea por observación, a partir de un ejemplar vivo, por la preservación de un espécimen o basándose en la literatura.

### C. Diseño de las consultas multidimensionales

Mediante las consultas multidimensionales, podemos realizar cálculos sobre la información contenida en la bodega de datos, y visualizarla desde diferentes perspectivas (fig. 5). Los cálculos estarán relacionados con la métrica definida, mientras que las perspectivas dependerán de las dimensiones incluidas en el modelo, pudiendo hacer diferentes combinaciones entre dimensiones para obtener nuevas perspectivas de la información, y moverse además entre las jerarquías construidas para obtener diferentes niveles de detalle [6].

Para construir estas consultas se empleó el lenguaje de expresiones multidimensionales o MultiDimensional eXpressions (MDX, por sus siglas en inglés) de SQL Server Analysis Services. En la fig. 6 podemos apreciar el resultado de una de las consultas diseñadas, en donde el cálculo de las métricas aparece representado en forma de una matriz.

Las consultas creadas de esta forma se utilizaron para el desarrollo de la aplicación móvil.

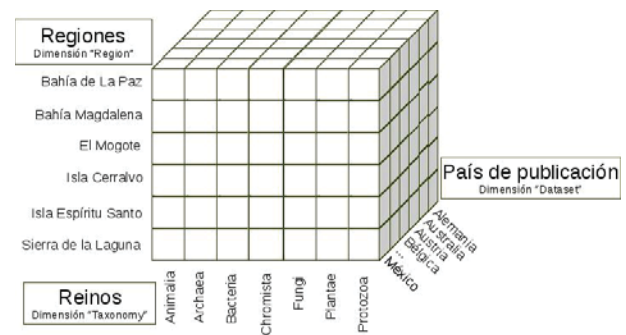


Fig. 5. Representación gráfica del cubo que muestra las diferentes dimensiones del modelo. En este aparecen señaladas las dimensiones “Region”, “Taxonomy” (en su nivel reino) y “Dataset” (en su nivel país de publicación).

	Animalia	Archaea	Bacteria	Chromista	Fungi	Plantae	Protozoa
Bahía de La Paz	8337	(null)	(null)	640	2	2008	176
Bahía Magdalena	5086	(null)	(null)	681	151	455	323
El Mogote	370	(null)	(null)	40	(null)	479	3
Isla Cerralvo	2421	(null)	(null)	(null)	(null)	1020	(null)
Isla Espíritu Santo	4969	(null)	(null)	(null)	(null)	1293	(null)
Sierra de la Laguna	12303	(null)	(null)	(null)	812	6437	(null)

Fig. 6. Resultado obtenido de la consulta MDX para reinos y regiones, en donde las etiquetas verticales a la izquierda representan los valores del atributo [Region Name] de la dimensión “Region” y las etiquetas horizontales sobre la matriz representan los valores del atributo [Kingdom] de la dimensión “Taxonomy”.

### D. Desarrollo de la aplicación móvil

La aplicación desarrollada se programó en el lenguaje Java, empleando un emulador de dispositivos móviles Android para su prueba. Dentro de su desarrollo se incluyó un mapa interactivo para consultar la información de las apariciones de organismos y las especies en una región geográfica, explotando los servicios web de GBIF y EOL. Además, se aplicó contenido algorítmico para mejorar la velocidad de respuesta de la aplicación, utilizando estructuras de listas ligadas y búsqueda binaria para hacer más rápidas las consultas de la información en memoria [9]. Junto a estos algoritmos, se empleó cómputo en paralelo a través de múltiples hilos de ejecución, con el objetivo de mejorar el rendimiento de la aplicación.

La apariencia de este mapa interactivo y la forma de consultar la información, se muestra en la fig. 7.

### E. Conexión del cliente móvil con el hiper cubo

Para que el cliente móvil pudiera acceder a la información del hiper cubo, se creó un servicio web basado en la tecnología ASP.Net MVC, utilizando el formato JSON para el intercambio de los datos. Este servicio puede ser explotado dentro de la aplicación móvil por medio de una interfaz gráfica, en la que el usuario selecciona las dimensiones que desea analizar, pudiendo tener hasta tres dimensiones en cada consulta (dos para los ejes  $x$  y  $y$  de las gráficas, y una tercera que funciona como filtro), para finalmente visualizar los datos por medio de tablas y distintos tipos de gráficos (ver fig. 8).

## III. RESULTADOS

Gracias a los servicios web explotados, la aplicación móvil brinda al usuario la información recolectada sobre los



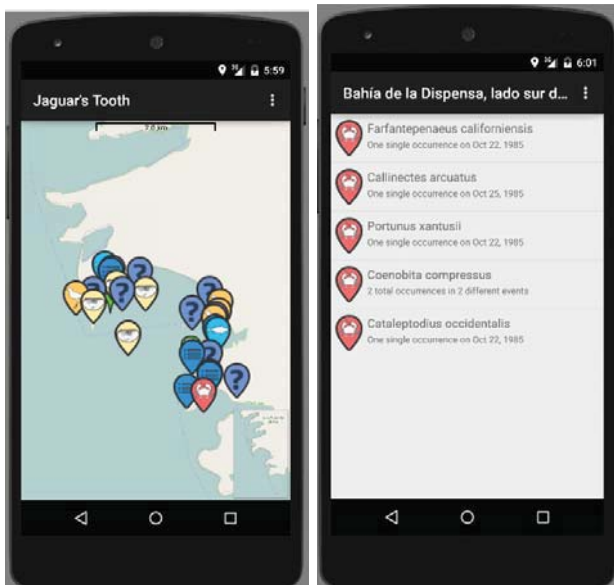


Fig. 7. Captura de la interfaz de la aplicación móvil, que muestra el mapa interactivo para consultar las apariciones de ejemplares en una región.

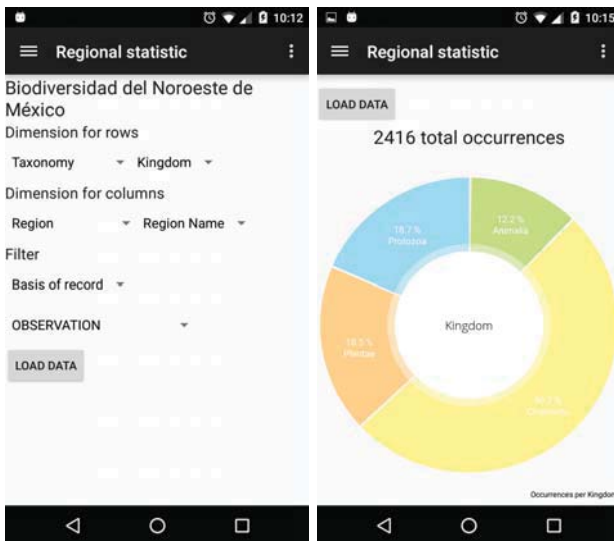


Fig. 8. Captura de la interfaz de la aplicación móvil, que muestra la apariencia del cliente para consultas multidimensionales y el resultado obtenido de una de estas consultas.

especímenes dentro de un mapa interactivo, que le permite acceder a la información biológica complementaria de las distintas especies encontradas en cada lugar, almacenando estos datos en la memoria interna del dispositivo, para consultarlos fuera de línea. Además, el modelo de datos diseñado, permite obtener información resumida sobre la biodiversidad de las regiones geográficas creadas, y utilizando la función de consulta del cliente móvil, podemos obtener una representación gráfica de esta información.

#### IV. CONCLUSIONES

Se logró comprobar la capacidad de las aplicaciones móviles para utilizarse como herramientas para la investigación en campo, funcionando como una fuente de información

primaria para el estudio de la diversidad de seres vivos sobre una región particular, fuera del alcance de las redes Internet. Además, se comprobó el potencial de los modelos de datos multidimensionales, para analizar la información histórica sobre la biodiversidad en distintas regiones geográficas de interés, para investigadores en las áreas de biología y ecología.

#### V. TRABAJO FUTURO

Este modelo multidimensional se diseñó de forma que puede ser extrapolado para cubrir un territorio mucho más grande, de modo que pueda analizarse una base de datos de biodiversidad extraída de GBIF que abarque el territorio mexicano completo, incluyendo su zona marítima, para de esta forma realizar un análisis mucho más profundo de la distribución de especies en las diferentes regiones de nuestro país, y utilizar la aplicación móvil para el trabajo de campo en las mismas.

#### AGRADECIMIENTOS

Agradezco su apoyo al Instituto Tecnológico de La Paz por brindarme la oportunidad de realizar este proyecto, así como al CONACyT por proporcionar los medios para llevarlo a cabo.

#### REFERENCIAS

- [1] Michael R. W. Rands, William M. Adams, Leon Bennun, Stuart H. M. Butchart, Andrew Clements, David Coomes, Abigail Entwistle, Ian Hodge, Valerie Kapos, Jörn P. W. Scharlemann, William J. Sutherland, and Bhaskar Vira. Biodiversity conservation: Challenges beyond 2010. *Science*, 329(5997):1298–1303, October 2010.
- [2] Meredith A. Lane Lane, L. Edwards James, Wouter Los, Cess H. J. Hof, Walter G. Berendsohn, Marc Geoffroy, Malcolm J. Scoble, Norman MacLeod, M. O'Neill, Steven A. Walsh, Gordon B. Curry, Richard J. Connor, Andrew C. Jones, Dagmar Triebel, Derek Persoh, Thomas H. Nash, Luciana Zedda, Gerhard Rambold, Richard J. White, Jacob Andersen Sterling, Ole Seberg, Chris J. Humphries, F. Borchsenius, and J. Dransfield. *Biodiversity Databases*. Systematics Association Special Volumes. CRC Press, first edition, April 2007.
- [3] Javier Otegui, Arturo H. Ariño, María A. Encinas, and Francisco Pando. Assessing the primary data hosted by the spanish node of the global biodiversity information facility (GBIF). *PLoS ONE*, 8(1):1–15, January 2013.
- [4] Michelle Rucker. Encyclopedia of life. *Reference Reviews*, 28(1):29–30, January 2014.
- [5] John L. Schnase, Judy Cushing, and James A. Smith. Biodiversity and ecosystem informatics. *Journal of Intelligent Information Systems*, 29(1):1–6, August 2007.
- [6] Petteri Vihervaara, Mia Ronka, and Mari Walls. Trends in ecosystem service research: Early steps and current drivers. *Ambio*, 39(4):314–324, June 2010.
- [7] Linda L. Hill. *Georeferencing: The Geographic Associations of Information*. MIT Press, 2009.
- [8] Alex Hardisty, Dave Roberts, and “The Biodiversity Informatics Community”. A decadal view of biodiversity informatics: challenges and priorities. *BMC Ecology*, 13(1):16, April 2013.
- [9] Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*. The MIT Press, 2009.

# Análisis del Patrón de Comportamiento en la Participación Ciudadana de Elecciones Estatales

Aplicación de Técnicas de Minería de Datos para el Análisis de Patrones

**M.S.C. José Luis Gutiérrez Villanueva**

Maestría en Sistemas Computacionales  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
jlgutierrezv87@gmail.com

**M.A.T.I. Luis Armando Cárdenas Florido**

División de Estudios de Posgrado e Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
armando.cardenas@gmail.com

**Abstract**— En una democracia, la participación ciudadana en los procesos electorales representa un aspecto muy relevante en la definición del futuro de un país. En este trabajo, a través de la aplicación de técnicas de minería de datos, se define el patrón de comportamiento detectado mediante el análisis de la participación ciudadana presentada en los procesos electorales realizados por el Instituto Estatal Electoral de Baja California Sur, en los cargos a elección de Gobernador del Estado,

A través de la detección de patrones de comportamiento, se busca apoyar en la definición de estrategias específicas y mejor dirigidas que permitan realizar una planeación mucho más eficiente de la promoción del voto en futuros procesos electorales.

**Palabras Clave**—Minería de Datos; Análisis de Patrones; Modelo de comportamiento; Elecciones; Resultados Electorales; Bodega de Datos.

## I. INTRODUCCIÓN

Actualmente vivimos en un mundo donde las nuevas tecnologías crecen a pasos agigantados y en intervalos de tiempo muy cortos. Tal es el caso especial del manejo y almacenamiento de la información mediante tecnologías computacionales, donde cada vez existe un mayor volumen de información y por lo tanto se deben generar nuevas técnicas de almacenamiento y procesamiento de la información.

El uso de bodegas de datos permite recopilar y almacenar grandes volúmenes de información de diferentes orígenes de datos, para ser analizada y procesada mediante diferentes técnicas, como por ejemplo el uso de Minería de Datos.[1]

Actualmente en el Estado de Baja California Sur se aplica un sistema democrático para la elección de Gobernador Constitucional, llamado Proceso Electoral, el cual es organizado por el Instituto Estatal Electoral de Baja California Sur (IEEBCS).

Dentro de las atribuciones del IEEBCS se encuentra una muy importante: la Promoción del Voto. Su principal objetivo es la de incentivar a los ciudadanos a que acudan a las urnas el día de la jornada electoral para ejercer su derecho a voto.

Cada vez que se realizan elecciones electorales se genera una gran cantidad de información valiosa. Una de ellas, y la más importante, son los resultados electorales. Estos son el fundamento y/o respaldo de la constancia de mayoría que se le otorga al candidato ganador de la elección.

La información que puede extraerse de los resultados electorales es muy valiosa y de mucha utilidad para los organismos electorales, ya que en base al análisis de los datos contenidos se puede realizar la planeación de la promoción del voto para procesos electorales próximos.

Es por ello la necesidad de un modelo de minería de datos para el análisis de la participación ciudadana en procesos electorales del estado de Baja California Sur, que lleve a la determinación de las variables que intervienen en sus patrones de comportamiento.

## II. METODOLOGÍA

Para poder realizar el modelado, se utilizó la metodología CRISP-DM, la cual es de las metodologías más utilizadas para la elaboración de proyectos de minería de datos. Está basada en actividades ordenadas en seis fases que recorren todo el proceso de minería de datos; desde la definición de los objetivos del negocio que se pretende obtener, hasta la vigilancia y el mantenimiento del modelo que se proponga e implemente.[2]

Cada una de las fases se subdivide a su vez en tareas ordenadas en un esquema jerárquico. Es una de las más completas y constituye un mapa de ruta que permite determinar qué actividades desarrollar en cada una de sus etapas. Contiene las fases de un proyecto, sus tareas respectivas, y las relaciones entre estas tareas.

Además algunas de las fases son bidireccionales, lo que significa que algunas fases permitirán revisar parcial o totalmente las fases anteriores.



Una de las ventajas de la metodología CRISP-DM es que este sistema está diseñado como una metodología independiente de la herramienta tecnológica a utilizar en la explotación de datos. Esto último la hace más flexible. Otro punto importante es que es de libre acceso y cumple con la característica de ser orientada al negocio.[3]

## APLICACIÓN DE LA METODOLOGÍA.

### A. *Comprensión del Negocio.*

#### 1.- Determinación de los objetivos del Negocio.

##### Contexto

La Dirección Ejecutiva de Organización Electoral es la encargada de llevar a cabo el proceso de ubicación de casillas, el cual consiste en la planeación e instalación de las casillas a utilizar en la Jornada Electoral. En base a este proceso se genera una base de datos con los datos técnicos y de ubicación de cada una de las casillas a instalar durante el proceso electoral.

Dentro de las funciones y actividades más importantes de la Dirección Ejecutiva de Educación Cívica y Capacitación Electoral se encuentra la Promoción del Voto, que se lleva a cabo mediante la creación de programas de promoción y difusión de la imagen institucional y el incentivo de la participación ciudadana.

En general una de las actividades principales realizadas por el Instituto y que se lleva previamente a un proceso electoral, es la Promoción del Voto. Esto consiste en la creación y planeación de estrategias y programas para promover la participación ciudadana, es decir, fomentar entre la población el acudir a ejercer su voto el día de la elección.

##### Objetivos

- La coordinación, preparación, desarrollo y vigilancia de los procesos electorales.
- Velar por la autenticidad y efectividad del sufragio.
- Llevar a cabo la promoción del voto y la difusión de la educación cívica y la cultura democrática.
- Todas las actividades del Instituto se regirán por los principios de certeza, legalidad, independencia, imparcialidad, máxima publicidad y objetividad.

El principal objetivo de esta investigación es poder identificar patrones de comportamiento mediante el análisis de la participación ciudadana presentada en procesos electorales pasados, organizados por el IEEBCS.

##### Criterios de éxito

Obtener un alto porcentaje en la participación ciudadana de procesos electorales futuros realizados por el IEEBCS, para lo que se requiere llevar a cabo una promoción eficaz del voto. También se desea poder realizar una promoción del voto más específica y enfocada con base a los diferentes sectores sociales de la ciudadanía, ya que actualmente se realiza una promoción del voto muy general.

Por lo tanto si mediante el análisis de los datos se pueden identificar los patrones de comportamiento en la participación ciudadana y con ello poder llevar a cabo una promoción del voto más específica por zonas, tendríamos un alto porcentaje de participación ciudadana en procesos futuros.

#### 2.- Evaluación de la Situación

##### Inventario de recursos

El IEEBCS ha proporcionado el apoyo necesario para llevar a cabo la investigación. Además de contar con las herramientas tecnológicas de hardware y software, colabora con el capital humano necesario para acceder y analizar los datos.

Se contó con el apoyo de la Unidad de Cómputo y Servicios Informáticos del IEE, integrada por un Director de la Unidad, un Coordinador de la Unidad, un Desarrollador de Sistemas y un Ingeniero de Soporte.

En la parte de la infraestructura, el Instituto cuenta tanto con servidores físicos como servidores virtuales. Los físicos como servidor de datos (HP ProLiant ML350p GenB), servidor de dominio (ProLiant ML110) y servidor proxy/firewall (HP DX2300). Los virtuales se utilizan como servidores web.

En la parte de software, se cuenta con Microsoft Windows Server 2012, Microsoft Windows SQL Server 2012, Microsoft Visual Studio 2012 Ultimate y Profesional.

##### Factibilidad

El costo beneficio para llevar a cabo la investigación ofreció un mayor grado de beneficio, ya que no se requirió un gran despliegue de recursos financieros, la mayoría de los requerimientos necesarios ya se cuentan con ellos.

Por lo tanto la implementación de este proyecto fue completamente factible considerando que los beneficios son mayores y la inversión necesaria para implementarlo es mucho menor, casi nulos.

### 3.- Determinación de los Objetivos de Minería de Datos

#### Objetivo de minería de datos

El objetivo principal por el cual se desea implementar un modelo de minería de datos para el análisis de la participación ciudadana es poder identificar los factores o características similares que existen entre las personas que acudieron a votar y también las que no ejercieron tal derecho.

Esto con la finalidad de identificar si existe un patrón similar entre los votantes y con base a ello realizar una promoción del voto basado en esos factores identificados en el patrón obtenido.

#### Criterios de éxito del modelo de minería de datos

Para poder decir que el análisis de la información mediante minería de datos fue exitoso en nuestro proyecto, debemos obtener los patrones de comportamiento en la participación ciudadana y poder identificar los factores o variables que influyen en el comportamiento, para llevar a cabo una planeación de la promoción del voto mucho más estratégica y especializada.

### 4.- Producir el Plan de Proyecto

#### Evaluación inicial de funciones y algoritmos

Se utilizó la herramienta de Microsoft SQL Server 2012, la cual cuenta con un módulo dedicado a la Inteligencia de Negocios, que permite la construcción de bodegas de datos y el uso de algoritmos de minería de datos, así como la generación de cubos y vistas para modelos multidimensionales.

Teniendo en cuenta las características de la información disponible y los objetivos definidos para el análisis y búsqueda patrones que se desean identificar, se seleccionaron tres tipos de algoritmos con los que se experimentaron y se realizaron pruebas preliminares, se busca descubrir hechos que ocurren en común dentro de un determinado conjunto de datos e identificar relaciones entre ellos, además de poder modelar atributos discretos y continuos.

Los algoritmos utilizados fueron:

- Algoritmo de Clúster
- Algoritmo de Árbol de Decisión

#### B. Comprensión de los Datos.

##### 1.- Recolección de Datos Iniciales

La información utilizada se encuentra distribuida en diferentes bases de datos, de las cuales en su mayoría son conformadas por el IEEBCS, por ejemplo la información de los resultados electorales de las elecciones llevadas a cabo en los años 2005 y 2011. Además se cuenta con una base de datos que contiene la ubicación de las casillas electorales en los mismos procesos, 2005 y 2011.

También se requirió de información de orígenes externos, como por ejemplo datos estadísticos y demográficos, proporcionados por el Instituto Nacional de Estadística y Geografía.



PATERNO	PEREZ	HERNANDEZ
MATERNO	LOPEZ	LEON
NOMBRE	JUAN	MARIA
EDAD	35	23
SEXO	H	M
CALLE	NAVEGANTES #45	PASEO DEL CORTEZ
COLONIA	CENTRO	PEDREGAL
ENTIDAD	3	3
MUNICIPIO	1	1
DISTRITO	1	1
SECCIÓN	121	122
CASILLA	B	C
CLAVE_ELEC	PRLPJU75012303H500	HRLNMA87020303M200
VOTÓ	SI	NO

##### 2.- Reporte de Recolección de Datos Iniciales

Una vez llevada a cabo la recolección de los datos iniciales, correspondientes a los resultados electorales de los procesos llevados a cabo en 2005 y 2011, obtuvimos la siguiente información. A manera de ejemplo ilustrativo solo se muestran datos del proceso 2011:

###### • Ubica Tu Casilla 2011

Campo	Datos de Ejemplo	
ENTIDAD	3	3
MUNICIPIO	1	1
DISTRITO	1	1
SECCIÓN	135	136
CASILLA	B	C
TIPO_DOMICILIO	PARTICULAR	ESCUELA
DOMICILIO	GUILLERMO PRIETO #987	ESC. PRIMARIA 3 DE MAYO
MANZANA	78	23

###### • Participación Ciudadana 2011

Campo	Datos de Ejemplo	
PATERNO	PEREZ	HERNANDEZ
MATERNO	LOPEZ	LEON
NOMBRE	JUAN	MARIA
EDAD	35	23
SEXO	H	M
CALLE	NAVEGANTES #45	PASEO DEL CORTEZ
COLONIA	CENTRO	PEDREGAL
ENTIDAD	3	3
MUNICIPIO	1	1
DISTRITO	1	1
SECCIÓN	121	122
CASILLA	B	C
CLAVE_ELEC	PRLPJU75012303H500	HRLNMA87020303M200
VOTO	SI	NO

#### C. Preparación de los Datos.

##### 1.- Seleccionar y Verificar los Datos

Con base a la recolección de los datos realizada con anterioridad y su análisis, procedimos a seleccionar los atributos útiles y de mayor relevancia de nuestros datos, así

como la exclusión de aquellos que no fueron necesarios para cumplir nuestros objetivos de minería de datos.

Campo	Estatus
PATERNO	EXCLUIDO
MATERNO	EXCLUIDO
NOMBRE	EXCLUIDO
EDAD	INCLUIDO
SEXO	INCLUIDO
CALLE	EXCLUIDO
COLONIA	INCLUIDO
ENTIDAD	EXCLUIDO
MUNICIPIO	INCLUIDO
DISTRITO	INCLUIDO
SECCION	INCLUIDO
CASILLA	INCLUIDO
CLAVE_ELEC	EXCLUIDO
VOTO	INCLUIDO

### 3.- Integración de los Datos

En primer lugar se realizó la integración de los datos entre la tabla de Participación Ciudadana y la tabla Ubica Tu Casilla. La integración fue posible mediante los campos en común: “sección” y “casilla”. De este modo pudimos integrar los datos del tipo de casilla y la manzana a la que pertenece.

#### Participación Ciudadana

Campo	Datos de Ejemplo	
EDAD	35	23
SEXO	H	M
COLONIA	CENTRO	PEDREGAL
MUNICIPIO	1	1
DISTRITO	1	1
SECCIÓN	121	122
CASILLA	B	C
VOTO	SI	NO

#### Ubica Tu Casilla

Campo	Datos de Ejemplo	
SECCION	123	124
CASILLA	B	C
TIPO_DOMICILIO	PARTICULAR	ESCUELA
MANZANA	23	56

Posteriormente se llevó a cabo la integración de la tabla generada en la primera integración con el tercer origen de datos seleccionado para nuestro modelo, el cual se refiere a los datos geo electorales. En este caso en particular la integración se realizó mediante el campo común “sección”.

### Tabla Final después de la integración.

Campo	Datos de Ejemplo	
SECCION	121	122
CASILLA	B	C
TIPO_DOMICILIO	PARTICULAR	ESCUELA
MANZANA	23	56
EDAD	35	23
SEXO	H	M
COLONIA	CENTRO	PEDREGAL
MUNICIPIO	1	1
DISTRITO	1	1
VOTO	SI	NO
GRA_PROM_ESC	10	5
18YMAS_F	307	505
18YMAS_M	306	905
18YMAS_EDUC_PB	340	185
CASADOS	379	360
VIVHAB_MAT	248	181
VIVHAB_SINMAT	2	29
VIVHAB_PC	117	67
VIVHAB_TV	246	205
VIVHAB_AUT	195	162

### 4.- Proceso de Extracción, Transformación y Carga

Los campos que fueron sometidos al proceso de transformación fueron: escolaridad, posbasica, casados, vivpc, vivau y se agregó un nuevo campo llamado edadr.

- **ESCOLARIDAD.-** En este campo se hizo una transformación de los valores para clasificarlos por Primaria (P1), Secundaria (S), Preparatoria (P2) y Universidad o posteriores (U).

#### Condiciones utilizadas en la transformación del campo ESCOLARIDAD.

```
IF escolaridad<7
    escolaridad='P1'
ELSE IF escolaridad<10
    escolaridad='S'
ELSE IF escolaridad<13
    escolaridad='P2'
ELSE
    escolaridad='U'
```

- **POSBASICA/CASADOS/VIVPC/VIVAU.-** En el caso de estos campos se hizo una transformación de los valores para clasificarlos por Casi ninguno (CN), Menos de la mitad (ME), Mas de la mitad (MA) y Casi todos (CT).

### Condiciones utilizadas en la transformación de los campos POSBASICA, CASADOS, VIVPC Y VIVAU.

```

IF posbasica < 25
    posbasica = 'CN'
ELSE IF posbasica < 50
    posbasica = 'ME'
ELSE IF posbasica < 75
    posbasica = 'MA'
ELSE
    posbasica = 'CT'

```

- EDADR.- Este campo se va a generar en base a la edad de las personas, el cual nos va a permitir clasificar a las mismas por rango de edad. Los rangos creados son: Joven (JO), Adulto (AD) y Tercera Edad (3A).

### Condiciones utilizadas en la transformación del campo EDADR.

```

IF edad < 30
    edadr = 'JO'
ELSE IF edad < 60
    edadr = 'AD'
ELSE
    edadr = '3E'

```

## 5.-Modelo Multidimensional

El modelo que se creó para nuestro objeto de estudio es un modelo multidimensional de tipo Copo de Nieve. El cual cuenta con 3 dimensiones principales, las cuales son: Personas, Sección y Población.



La dimensión Personas cuenta con los datos personales de los ciudadanos que acudieron a votar y los que no. Por ejemplo, si votó, la edad, sexo, rango de edad y escolaridad. A su vez esta dimensión cuenta con las dimensiones Edad, que muestra la edad de las personas y su rango de edad al que pertenecen los ciudadanos. De la dimensión Edad se desprende la dimensión de RangoEdad, que muestra el rango de edad a la que pertenecen las personas, Jóvenes (JO), Adultos (AD) y 3ra Edad (3E). Además cuenta con una dimensión Escolaridad, la cual se refiere al grado de estudios de las personas, Primaria (P1), Secundaria (S), Preparatoria (P2) y Universidad o posteriores (U). Estas dimensiones son necesarias ya que con base a este modelo de dimensiones, podemos visualizar

información de diferentes aspectos, por ejemplo las edades de las personas que votaron. También podemos consultar el número de personas que votaron por su rango de edad, y con ello identificar el rango de edad que más voto en una elección. Incluso si queremos un nivel más de detalle, por ejemplo, saber la edad de las personas que más votaron.

La dimensión Sección nos permite agrupar a los ciudadanos por distintas secciones electorales, distrito al que pertenecen y el municipio al que pertenece el distrito y las secciones que conforman un distrito. También cuenta con las dimensiones Distrito, la cual nos permite mostrar a los ciudadanos de un distrito y poder obtener información como por ejemplo el distrito que más votó y a su vez identificar cual sección fue la que más voto de ese mismo distrito. De igual forma cuenta con la dimensión Municipio, con la que podemos mostrar los ciudadanos que votaron por municipio, identificar qué municipio fue el que más votó y de ese municipio, cual fue el distrito que más votó.

La dimensión Población agrupa la información de datos estadísticos poblacionales de los ciudadanos que acudieron a votar. Cuenta con la dimensión Auto, la cual nos permite conocer si las personas que votan cuentan con al menos un automóvil. La dimensión PC nos permite conocer si las personas cuentan con un equipo de cómputo. La dimensión TV nos muestra información relativa a si las personas cuentan con una televisión en su vivienda.

## D. Modelado.

### 1.- Técnicas de Modelado

#### a).- Algoritmo de Clasificación

Es un algoritmo que permite clasificar los datos con base a ciertos criterios y atributos. Nos permite identificar los atributos de mayor relevancia o influencia sobre algún campo en particular que previamente seleccionamos para su análisis.[4]

Para nuestro modelo en particular se utilizó el método de clasificación Árboles de Decisión.

#### Árbol de Decisión

Este método hace predicciones basándose en las relaciones entre las columnas de entrada de un conjunto de datos. Utiliza los valores de estas columnas para predecir los estados de una columna que se designa como elemento de predicción. Específicamente, identifica las columnas de entrada que se correlacionan con la columna de predicción.

Aplicamos este método para identificar los campos que se relacionan con el campo seleccionado para predicción: VOTÓ, es decir, los valores o parámetros que influyen en que la persona vote o no. En pocas palabras, los factores tanto personales como sociales que influyen en la participación ciudadana.

### b).- Algoritmo de Clúster

Es un algoritmo que permite la segmentación de los datos con base a ciertos criterios y poder obtener datos similares o con características parecidas y con base a ello generar clústeres o grupos de datos. Utiliza técnicas iterativas para agrupar los casos de un conjunto de datos dentro de clústeres que contienen características similares.[5]

Para nuestro modelo en particular se utilizó el método de agrupamiento K-Means.

#### K-Means

Es un método de agrupamiento que tiene como objetivo la partición de un conjunto de N observaciones en K grupos, en el que cada observación pertenece al grupo más cercano a la media. Tiene por objeto encontrar grupos de observaciones (clústeres) con características semejantes. Las observaciones de un grupo deben ser parecidas a las otras observaciones del grupo, pero deben ser lo más diferentes de las observaciones de otros grupos.

Este modelo permite identificar o agrupar a las personas que votaron y que no votaron y que cuentan con características similares. Estas características a identificar serán las columnas de entrada que seleccionaremos, y con base a este agrupamiento, poder definir un patrón o patrones de comportamiento en la participación ciudadana.

### 2.- Diseño de Pruebas

Para poder llevar a cabo el análisis de la información y obtener resultados en menor tiempo posible, toda vez que el universo de los datos es muy extenso (hablamos de una bodega de datos con 960,019 registros), se tomaron como objeto de estudio solo los datos correspondientes al municipio de La Paz (73,901 registros). Además, cada municipio es muy diferente uno de otro en cuanto a su nivel de participación ciudadana, así como sus datos estadísticos.

### 3.- Construcción del Modelo

Una vez que seleccionamos nuestra técnica de modelado y diseñamos las pruebas, procedemos a realizar la construcción de nuestro modelo.

Después se procede al análisis de la información mediante la implementación y ejecución de los algoritmos de minerías de datos seleccionados.

#### Árbol de Decisión

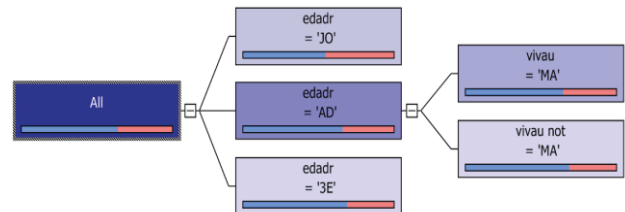
Primeramente analizamos nuestra información por medio del algoritmo de clasificación, Árbol de Decisión. Esto con la finalidad de detectar los valores más representativos o de mayor relevancia en las personas que acuden a votar. Es decir, los factores que influyen en las personas para ejercer su voto el día de la elección.

Después de varias ejecuciones del algoritmo, modificando los parámetros de entrada, pudimos obtener resultados útiles para nuestro estudio. Los parámetros establecidos en la última ejecución fueron los siguientes:

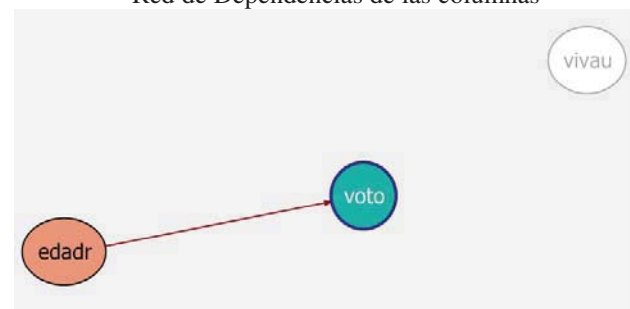
Columna de análisis: VOTO.

Columnas de entrada: EDADR, VIVPC y VIVAU.

#### Árbol de Decisiones Resultante



#### Red de Dependencias de las columnas



**Conclusiones:** Los resultados obtenidos en la ejecución del algoritmo Árbol de Decisión nos permiten establecer las variables de mayor influencia en la participación ciudadana así como su orden jerárquico. Es decir, la variable de mayor influencia en los votantes es la EDADR. Como segunda variable es VIVAU. También como tercera variable influyente es VIVPC.

#### K-Means

En esta etapa analizamos nuestra información por medio del algoritmo de clúster o agrupamiento, utilizando el método de agrupación K-Means, para agrupar los votantes con características similares. Con esto podremos identificar y/o establecer los patrones de comportamiento.

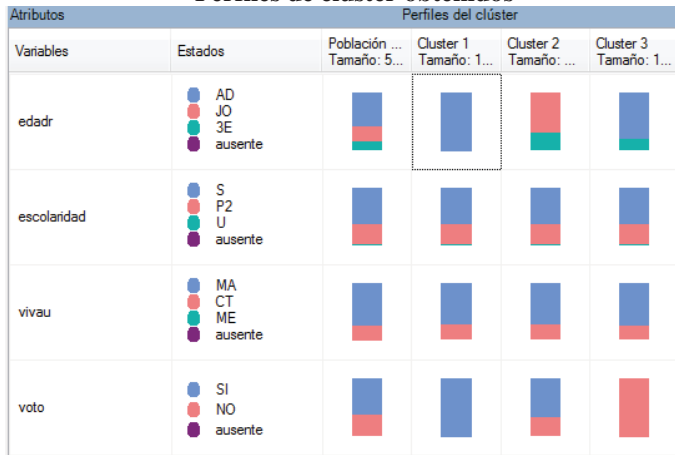
Después de varias ejecuciones del algoritmo, modificando los parámetros de entrada, pudimos obtener resultados útiles para nuestro estudio. Los parámetros establecidos en la última ejecución fueron los siguientes:

Número de Clúster: 3.

Método de Agrupamiento: K-Means Escalable.

Columnas de entrada: ESCOLARIDAD, VOTO, EDADR y VIVAU.

### Perfiles de clúster obtenidos



**Conclusiones:** Los resultados obtenidos en la ejecución de este algoritmo permiten establecer una tendencia en la participación ciudadana y los patrones de comportamiento de los ciudadanos.

### III. RESULTADOS OBTENIDOS

#### 1.- Variables de mayor relevancia en los votantes

Las variables más representativas en la participación ciudadana, ordenadas de mayor a menor nivel de influencia, son:

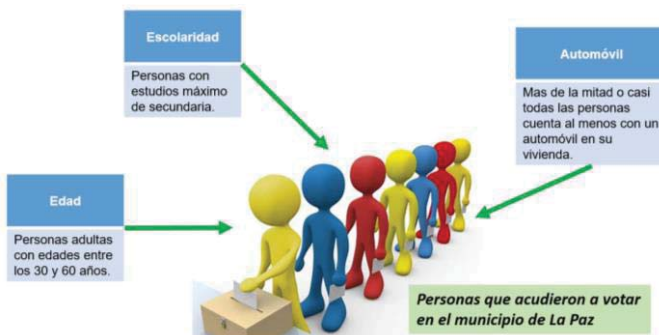
EDAD	Rango de edad de los votantes.
VIVAU	Vivienda con automóvil.
ESCOLARIDAD	Grados de escolaridad.

Como conclusión se establece que la variable de mayor influencia en la participación ciudadana es la edad del votante, y en segundo lugar si cuenta con automóvil en su vivienda. Con menor grado de influencia obtuvimos su grado de escolaridad.

#### 2.-Patrón de Comportamiento

El patrón de comportamiento detectado en el municipio de La Paz fue:

*“Las personas que acuden a votar son adultas, con edades entre los 30 y 60 años; con escolaridad máxima de secundaria y cuentan al menos con un automóvil en su vivienda.”*



### 3.- Impacto de los Resultados

Previo a la jornada electoral se lleva a cabo una actividad llamada “Promoción del Voto”, la cual consiste en motivar a la ciudadanía para que asistan a votar el día de las elecciones. Esta promoción se lleva a cabo de manera muy general, es decir, aplicando el mismo programa o método de promoción a todas las poblaciones del estado de Baja California Sur.

Sin embargo, en las ciudades y los poblados muchas veces se cuentan con características diferentes, por lo que la mayoría de las veces la promoción del voto no tiene el mismo impacto o éxito en los diferentes lugares donde se aplica.

Por lo tanto gracias al análisis de la participación ciudadana de procesos anteriores y la detección de patrones de comportamiento en las personas que acudieron a votar en el municipio de La Paz, se puede llevar a cabo una planeación estratégica de la promoción del voto para futuros procesos electorales, mucho más enfocada hacia las personas con características particulares de ese municipio.

### IV. TRABAJOS FUTUROS

- Planeación de una promoción del voto especializada para el municipio de La Paz, tomando en cuenta los resultados obtenidos en este trabajo.

- Aplicar el modelo de minería de datos para los demás municipios del Estado de Baja California Sur, y de este modo llevar a cabo una promoción del voto especializada con base al comportamiento de la participación ciudadana.

- Incluir más variables o métricas en el modelo establecido como datos de entrada. Por ejemplo, datos demográficos, culturales, sociales, censos de población.

- Aumentar la generación de bases de datos con mayor volumen de información, que puedan enriquecer el análisis de la información en la búsqueda de nuevos patrones en los próximos procesos electorales.

- Generación de nuevos modelos de minería de datos con objetivos distintos y el descubrimiento de nueva información y/o conocimiento en los procesos electorales.

### V. REFERENCIAS

- [1] Hernández Orallo, J. (2004). Introducción a la Minería de Datos: Pearson.
- [2] Dario, B. R. (2010). Metodología para la construcción de un data warehouse. Córdoba, Argentina: Hefesto.
- [3] Han, J.; Kamber M. (2001). Data Mining: Concepts and Techniques. Morgan Kaufmann Publishers, USA.
- [4] Petrovic, D. (2010). SQL Server 2008 Manual de referencia. Alemania: Mc Graw Hill.
- [5] Microsoft.com 2013, Database System, Performance and Scalability, SQL Server 2012 Business Intelligence Editions.



# Implementación de algoritmos de control en un péndulo de transmisión directa

O. Tepal Nieto, O. Félix Beltrán, F. Reyes Cortés, J. Cid Monjaraz

*Facultad de Ciencias de la Electrónica, Benemérita Universidad Autónoma de Puebla, Av. San Claudio y 18 Sur, C.U., Col. Jardines de San Manuel, Puebla, Pue., C.P.72570 México*  
(oswald.tepaln@alumno.buap.mx).

**Resumen--** En el presente trabajo se presenta la implementación de tres algoritmos de control en un péndulo de transmisión directa poniendo en uso un sistema de control de bajo costo compuesto de una tarjeta de adquisición de datos y control de la instrumentación de su electrónica, para el control de un servo de transmisión directa. Los algoritmos de control implementados son el PD, atan y tanh, con los cuales se obtienen resultados con  $\tilde{q} = 1.46, 0.44, 0.28$ , respectivamente.

**Palabras Clave--**Control de fuerza; Péndulo; Robot de Transmisión Directa; Arduino.

## I. INTRODUCCIÓN

La robótica es un área de investigación y desarrollo de sistemas mecánicos conocidos como robots [1]. En particular, un robot de transmisión directa se define como un brazo mecánico donde todas sus articulaciones son manejadas con servos de transmisión directa [2]; a su vez, la transmisión directa consiste en que el rotor del motor esté directamente acoplado al eslabón sin necesitar de sistemas de transmisión mecánicos, como engranes [2].

Un ROTRADI (Robot de transmisión directa) es una plataforma experimental que se mueve en su espacio tridimensional equipado con tarjetas electrónicas y programado para realizar aplicaciones de automatización. Estas plataformas poseen características que pueden competir en desempeño y versatilidad con las mejores marcas mundiales de robótica. No obstante su principal característica es el hecho de tener arquitectura abierta, esto significa que se puede evaluar experimentalmente cualquier estrategia de control, en contraste con los robots industriales con arquitectura cerrada que no permiten programar nuevos esquemas de control y quedan confinados a lo que permita realizar su sistema operativo [3].

Desde la creación de estas plataformas hasta la fecha, han sido implementadas diversas tarjetas de adquisición de datos, como lo son las tarjetas Tag1 [4] y la serie de tarjetas MFIO [5]. Debido a la complejidad de estos sistemas, resulta complicada la implementación de nuevos algoritmos de control

para aquellos que no están familiarizados con el sistema. Por lo que, como parte de su actualización, se propone el uso de una tarjeta Arduino, pues en contraste con las tarjetas previamente implementadas, esta tarjeta es programable en un ambiente amigable, de fácil entendimiento, pero sobre todo de fácil acceso. Tomando esto en cuenta se inicia el proceso de implementación para el sistema más simple disponible, en este caso un ROTRADI de un grado de libertad, esto es, un péndulo de transmisión directa.

Específicamente, para esta investigación se escogió la tarjeta Arduino Mega 2560, ya que esta cuenta con compatibilidad con MATLAB. Arduino Mega es una placa microcontrolador basada ATmega1280. Tiene 54 entradas/salidas digitales (de las cuales 14 proporcionan salida PWM), 16 entradas digitales, 4 UARTS (puertos serie por hardware), un cristal oscilador de 16MHz, conexión USB y conector ICSP. El ATmega1280 tiene 128KB de memoria flash para almacenar código (4KB son usados para el arranque del sistema (bootloader), 8 KB de memoria SRAM y 4KB de EEPROM [6].

## II. CARACTERÍSTICAS DE LA PLANTA

La plataforma de prueba consta principalmente de tres elementos:

- a) El robot péndulo, en donde se contempla el sistema mecánico a controlar y la interfaz electrónica que permitirá el control del robot.
- b) El controlador, el algoritmo de control que gobernará el comportamiento del robot.
- c) El sensor, el encoder que permitirá conocer el posicionamiento del robot.

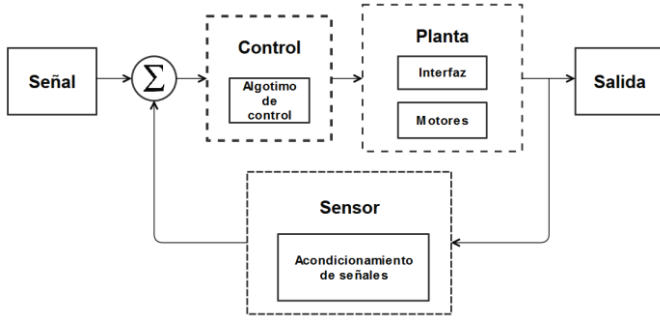


Figura 1 Diagrama a bloques del sistema de control.

Entre las características del robot se tiene que el actuador es un motor de transmisión directa DM1B-015G, cuyo par máximo es de 15 Nm y su encoder tiene una capacidad máxima de 2621440 pulsos por vuelta, lo que nos da una resolución de 0.00013 grados por pulso. Este motor es controlado por un driver DrvG3 el cual permite la programación de diversos parámetros de funcionamiento del motor dependiendo de la modalidad que se requiera, tales como control de posición, velocidad o torque.

Para la implementación abierta de algoritmos de control, se configura la planta en modo torque y adicionalmente se programa el driver para que el encoder funcione en modo diferencial con señales A-B en cuadratura desfasadas 90 grados, como se muestra en la Figura 2.

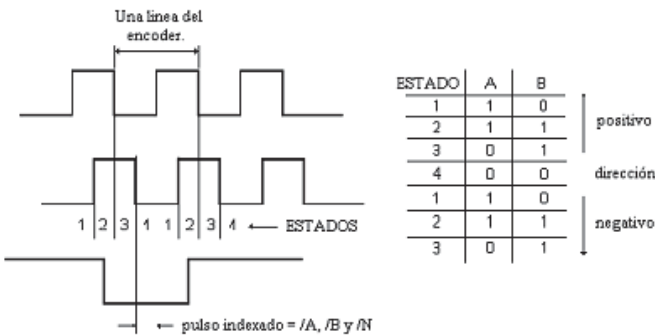


Figura 2 Señales del encoder de cuadratura A-B.

De igual forma y para propósitos de pruebas y ajustes, se configura el sistema para reducir la resolución del encoder a 10260 pulsos por vuelta que nos da una resolución aproximada de 0.035 grados por pulso.

### III. CINEMÁTICA

En esta sección se describe el modelo cinemático del robot de 1 grado de libertad (Péndulo) tomando como referencia la Figura 3. Para esto se usa el método de Euler-Lagrange, en donde se da por hecho que el origen del eje z está en el centro del rotor [5],

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} l_c \sin(q) \\ -l_c \cos(q) \end{bmatrix}, \quad (1)$$

donde  $l_c$  es la distancia al centro de masa y  $q$  es el ángulo.

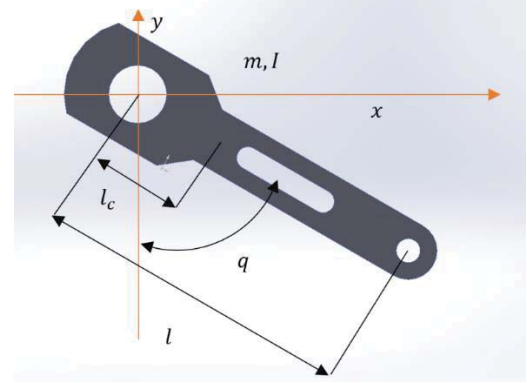


Figura 3 Esquema referencial del péndulo.

El modelo dinámico de un péndulo proporciona una descripción completa entre el par aplicado al servomotor y el movimiento de la estructura mecánica [3]. El lagrangiano está dado por:

$$\mathcal{L}(q, \dot{q}) = K(q, \dot{q}) - U(q), \quad (2)$$

$$\frac{1}{2} [ml_c^2 + I] \dot{q}^2 - mgl_c [1 - \cos(q)], \quad (3)$$

donde  $\dot{q}$  es la velocidad angular,  $m$  es la masa del eslabón,  $I$  es el momento de inercia del eslabón y  $g$  es la aceleración de la gravedad.

Las ecuaciones de movimiento de Euler-Lagrange para el caso particular del péndulo adquieren la forma de una ecuación escalar con la siguiente estructura [5]:

$$\tau = \frac{d}{dt} \left[ \frac{\partial \mathcal{L}(q, \dot{q})}{\partial \dot{q}} \right] - \left[ \frac{\partial \mathcal{L}(q, \dot{q})}{\partial q} \right] + f_f(f_e, \dot{q}) \quad (4)$$

donde  $\tau$  es el par aplicado y  $f_f$  es el fenómeno de fricción modelado por la ecuación siguiente:

$$f_f(f_e, \dot{q}) = b\dot{q} + f_c \text{signo}(\dot{q}) + f_e [1 - |\text{signo}(\dot{q})|], \quad (5)$$

donde  $f_c$  es la fricción de Coulomb,  $f_e$  es la fricción estática,  $b$  es el coeficiente de fricción viscosa y la función  $\text{signo}(\dot{q})$  está definida como:

$$\text{signo}(\dot{q}) = \begin{cases} 1 & \text{si } \dot{q} > 0, \\ -1 & \text{si } \dot{q} < 0. \end{cases} \quad (6)$$

Por tanto el modelo dinámico de un robot péndulo incluyendo el fenómeno de fricción está dado por:

$$\tau = [ml_c^2 + I] \ddot{q} - mgl_c \text{sen}(q) + b\dot{q} + f_c \text{signo}(\dot{q}) + f_e [1 - |\text{signo}(\dot{q})|], \quad (7)$$

donde  $\ddot{q}$  es la aceleración angular.

### IV. INTERFAZ ELECTRÓNICA

Para el control del sistema se propone el uso de una tarjeta Arduino Mega, destacándose por poseer una gran cantidad de librerías y soporte, además de su bajo costo y adecuadas prestaciones para los propósitos de cálculo y control. Para el control de la plataforma se precisa de la aplicación de voltaje, el cual debe estar en un rango de  $\pm 10V$  y cuyo valor va

a definir el par aplicado. Tomando esto en cuenta, la forma más apropiada para controlar el motor es a través de un DAC. Para esto se diseña un DAC incremental de 10 bits, tanto para valores de voltaje negativos como positivos como se muestra en las Figuras 4 y 5.

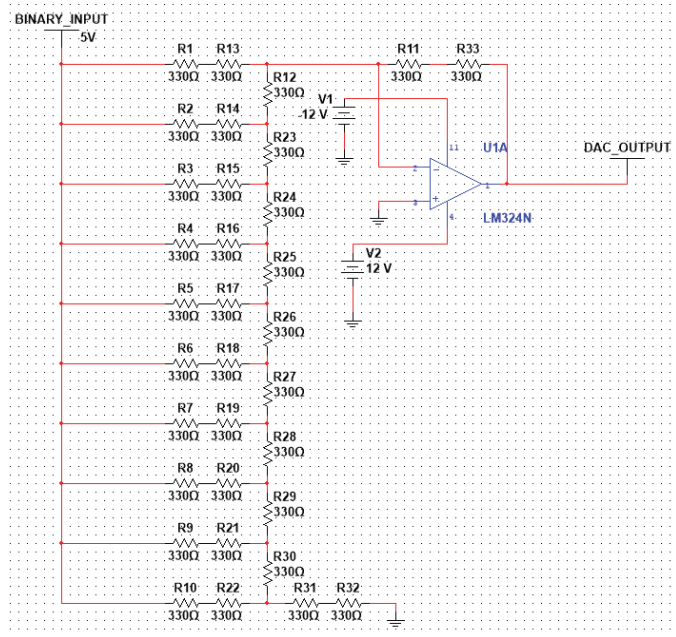


Figura 4 Circuito del DAC utilizado para el control del voltaje.

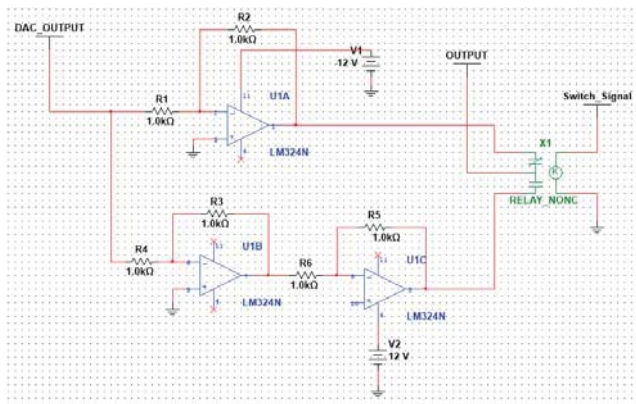


Figura 5 Circuito del DAC utilizado para el control del voltaje.

En la Figura 4 se muestra el típico circuito de un DAC incremental implementado con opamps y resistencias; las entradas al DAC provienen de las salidas digitales del Arduino, que en conjunto con el circuito devuelven un nivel de voltaje negativo, según el valor binario de 10 bits enviado por el microcontrolador. Tomando en cuenta que se requieren valores de voltaje negativos y positivos, se implementa un segundo circuito mostrado en la Figura 5, con el cual se invierte el voltaje y adicionalmente se da la posibilidad de controlar un onceavo bit dando 1023 valores de voltaje positivos y 1023 valores de voltaje negativos.

Para el control del sistema se diseñó un programa descrito en el diagrama de flujo de la Figura 6.

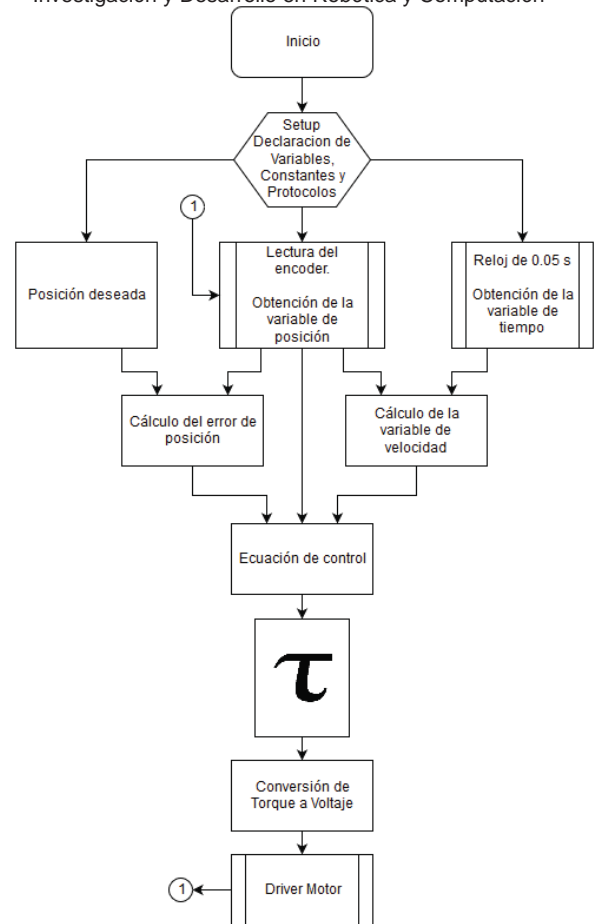


Figura 6 Diagrama de flujo del programa de control.

La función del programa se resume en la obtención de la posición actual del robot, a partir de la cual se obtienen parámetros como velocidad angular y el error de posición entre otros, con base en estos valores se genera un valor de par que es convertido a un equivalente en voltaje, para el caso específico del motor utilizado, el valor de  $\pm 10V$  corresponden a un par aplicado de  $\pm 15Nm$ .

Hay que resaltar que la principal forma de obtener el posicionamiento es a través de las rutinas de interrupción, las cuales son recursos o mecanismos del microcontrolador para responder a eventos, permitiendo suspender temporalmente el programa principal para ejecutar una subrutina de servicio de interrupción (ISR, por sus siglas en inglés) [6]. La tarjeta Arduino Mega tiene capacidad para ejecutar seis rutinas de interrupción a la vez y cada encoder requiere de dos rutinas de interrupción para realizar una correcta lectura de la posición; una vez terminada dicha subrutina se reanuda la ejecución del programa principal.

## V. ALGORITMOS DE CONTROL

Dada la estructuración del programa de control, resulta sencillo cambiar el algoritmo de control, pues todo esto se implementa en un simple bloque donde se hacen los cálculos necesarios para la obtención del torque necesario para mover el eslabón a la posición deseada.

Para este trabajo se consideran 3 algoritmos de control, los cuales se describen a continuación.

a) **Controlador PD más compensación de gravedad.**

Para propósitos de prueba y referencia, se implementa en primera instancia el controlador PD más compensación de gravedad, el cual está dado por la expresión [7] [8] [9]:

$$\tau = K_p \tilde{q} - K_v \dot{q} + g(q). \quad (8)$$

El controlador PD es ampliamente conocido y simple de implementar, los resultados de las pruebas con el controlador se pueden tomar como referencia para la comparación de desempeño de otros controladores y a su vez como un punto de partida para sistemas experimentales como el del sistema propuesto.

b) **Controlador atan-atan más compensación de gravedad.**

El segundo controlador propuesto se describe por [7] [8]:

$$\tau = K_p \operatorname{atan}(\tilde{q}) - K_v \operatorname{atan}(\dot{q}) + g(q). \quad (9)$$

c) **Controlador tanh-tanh más compensación de gravedad.**

Finalmente, el tercer controlador propuesto está dado por [7] [8]:

$$\tau = K_p \tanh(\tilde{q}) - K_v \tanh(\dot{q}) + g(q). \quad (10)$$

En los tres algoritmos de control,  $K_p$  es la constante de proporcionalidad,  $K_v$  es la constante derivativa y  $\tilde{q}$  es el error de posición.

VI. RESULTADOS

En esta sección se presentan los resultados obtenidos, las gráficas del par aplicado y de su posición angular con respecto al tiempo de cada uno de los controladores utilizados en la pruebas. Para cada prueba se seleccionó un valor en específico como ángulo deseado, en este caso ese valor corresponde a  $q_d = -90^\circ$  y el error de posición se calcula como:

$$\tilde{q} = q_d - q. \quad (11)$$

para las condiciones de estas pruebas, se realizaron 30 repeticiones.

A. Controlador PD

Para este controlador los resultados se pueden apreciar en la Figura 7, en donde el error de posición promedio fue de  $\tilde{q} = 1.46$  grados.

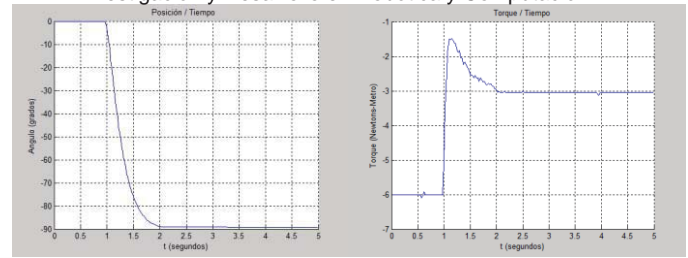


Figura 7 Gráfica de posición y torque con respecto al tiempo para el control PD.

B. Controlador atan-atan

Para este controlador los resultados se pueden apreciar en la Figura 8, en donde el error de posición promedio fue de  $\tilde{q} = 0.44$  grados.

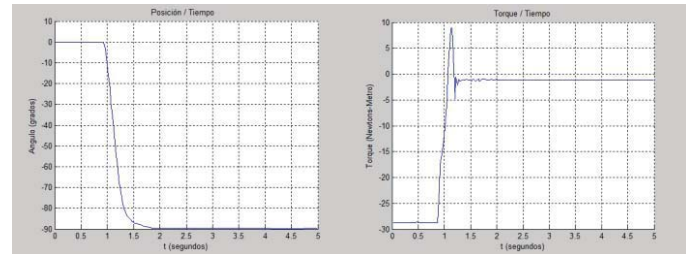


Figura 8 Gráfica de posición y torque con respecto al tiempo para el control atan-atan.

C. Controlador tanh-tanh

Para este controlador los resultados se pueden apreciar en la Figura 9, en donde el error de posición promedio fue de  $\tilde{q} = 0.28$  grados.

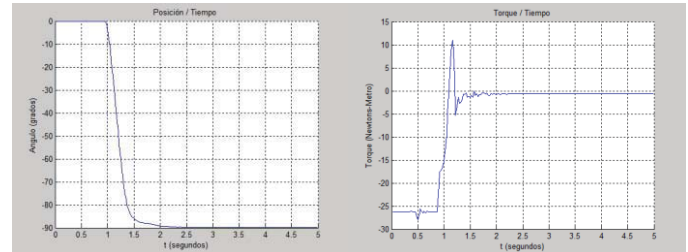


Figura 9 Gráfica de posición y torque con respecto al tiempo para el control tanh-tanh.

A continuación se muestra una tabla comparando el error de posición obtenida en cada uno de los controladores.

TABLA I. TABLA COMPARATIVA DEL PROMEDIO DEL ERROR.

Controlador	$\tilde{q}$ (grados)
PD	1.46
<i>tanh-tanh</i>	0.28
<i>atan-atan</i>	0.44

VII. CONCLUSIONES

Para este trabajo, el desempeño de los controladores a través del sistema propuesto fue eficaz con errores de

posicionamiento aceptables, cabe mencionar que este error puede ser reducido a través de una buena sintonización de las constantes  $k_p$  y  $k_v$ . Adicionalmente, el microcontrolador utilizado y el hardware diseñado para la planta, demostró ser versátil y completamente funcional para el procesamiento de información, por lo que se puede concluir que el Arduino es un sistema apto para la implementación de algoritmos experimentales de control y por tanto candidato para la implementación de la plataforma ROTRADI de tres grados de libertad en el cual se está trabajando.

La electrónica que se diseñada para el robot de 1 grado de libertad, tiene la cualidad de ser escalable, lo cual significa que se puede copiar e implementar para cualquier número de servomotores.

[1] L. Scavicco y B. Siciliano, *Modeling and Control of Robot Manipulador*, McGraw-Hill 2011.

[2] C. Chávez Olivares, F. Reyes Cortés, E. González Galván, M. Mendoza Gutierrez, I. Bonilla-Gutierrez. "Experimental evaluation of parameter identification schemes on an anthropomorphic direct drive robot". *International Journal of Advanced Robotic Systems*, pp. 9. ISO 690, 2012.

[3] F. Reyes, *Robótica, Control de Robots Manipuladores*, Alfa-omega, 1<sup>ra</sup> ed. 2011.

[4] E. Leon Bonilla, "Tarjeta PCI para controlar un Robot con tres grados de libertad", Tesis de Maestría en Ciencias de la electrónica opción Automatización, FCE-BUAP, 2007.

[5] A. Palma Asunción, "Control visual de robots manipuladores", Tesis de Maestría en Ciencias de la electrónica opción Automatización, FCE-BUAP, 2016.

[6] C. RUIYAN. "Principle and application of Arduino. *Electronic Design Engineering*", vol. 20, no. 16, pp. 155-157, 2012.

[7] J. G. Cebada Reyes, J. Tuxpan Meneses, P. Sánchez Sánchez, F. Reye Cortés, "Evaluación de 4 Estructuras de Control Mediante un Simulador Basado en un Robot de 2 Grados de Libertad", 8° Congreso Nacional de Mecatrónica, 2009.

[8] G. PALLI, C. MELCHIORRI. "Robust control of robots with variable joint stiffness". En *Advanced Robotics. ICAR 2009, International Conference on. IEEE*, pp. 1-6, 2009.

[9] O. R. Arroyo, F. R. Cortes, M. A. D. V. Trevino, S. V. Limon. "Variable Control Gains for PD Controller for Robot Manipulators". *Mechatronics, Electronics and Automotive Engineering (ICMEAE)*, pp. 187-192, 2015.

# Diseño de un controlador LQR para un giroscopio de 3 grados de libertad

Jerónimo Moyrón  
 Tecnológico Nacional de México  
 Instituto Tecnológico de La Paz,  
 Ingeniería Electromecánica  
 Blvd. Forjadores de B.C.S., La Paz, B.C.S., 23080,  
 MÉXICO  
 Email: moyron\_jero@hotmail.com

Jesús Sandoval  
 Tecnológico Nacional de México  
 Instituto Tecnológico de La Paz,  
 División de Estudios de Posgrado e Investigación  
 Blvd. Forjadores de B.C.S., La Paz, B.C.S., 23080,  
 MÉXICO  
 Email: jsandoval@itlp.edu.mx

Víctor Santibáñez  
 Tecnológico Nacional de México  
 Instituto Tecnológico de La Laguna,  
 División de Estudios de Posgrado e Investigación,  
 Apdo. Postal 49, Adm. 1, Torreón, Coahuila, 27001,  
 MÉXICO  
 Email: vsantiba@itlailaguna.edu.mx

**Abstract**—En este trabajo se presenta el diseño de un regulador cuadrático lineal para el control de un giroscopio de tres grados de libertad. La naturaleza subactuada del giroscopio es un desafío para su control, lo cual lo convierte en una excelente plataforma de prueba de diferentes algoritmos de control. Resultados experimentales son mostrados para validar el desempeño de nuestro sistema de control.

## I. INTRODUCCIÓN

El presente trabajo es una extensión del presentado en [1], en el cual se obtiene el modelo matemático de un giroscopio de tres grados de libertad, de la marca Quanser, con la finalidad de verificar el desempeño de diferentes algoritmos de control en simulación antes de su implementación. Para validar el modelo matemático en [1] se diseñó un compensador en adelante para el control de posición del giroscopio con resultados satisfactorios.

Debido a la naturaleza subactuada, ya que comúnmente el giroscopio tiene más grados de libertad que entradas de control, su estabilización se convierte en un desafío y contar con una plataforma didáctica para validar diferentes algoritmos resulta de interés para tópicos de investigación. En este trabajo presentamos el diseño de un regulador cuadrático lineal (LQR, por sus siglas en inglés) de un giroscopio Quanser, así como resultados experimentales que muestran un mejor desempeño con respecto al compensador presentado en [1]. Algunos trabajos relevantes sobre el control del giroscopio pueden ser encontrados en [2], [3] y [4].

El resto del trabajo está organizado como sigue. En la siguiente sección se describe el diseño de un LQR para el control de posición de un giroscopio. Luego, se presentan resultados experimentales para validar el desempeño de nuestro LQR. Finalmente, se dan algunas conclusiones.

## II. GIROSCOPIO DE TRES GRADOS DE LIBERTAD

### A. Descripción del prototipo

El giroscopio de la marca Quanser es mostrado en la Figura 1. Está compuesto de un disco metálico montado dentro de un marco interno de color azul, el cual a su vez está acoplado a un marco externo de plástico gris transparente. Ambos marcos están acoplados por medio de un par de resortes que dan estabilidad mecánica y ajuste a la estructura entera (ver [5], [6]). El giroscopio se coloca sobre un servomotor Quanser (SRV02) para formar en conjunto un sistema electromecánico de tres grados de libertad (g.d.l.). A dicho sistema se le puede añadir mayor dificultad para el control al colocarlo sobre una base móvil, que permite que el sistema realice dos rotaciones independientes sobre el mismo eje vertical (el marco exterior y la base móvil mostrados en la Figura 1). Al considerar constante la velocidad angular del disco del giroscopio se reduce a un sistema de 2 g.d.l. (las cuales corresponden a la inclinación del disco y al desplazamiento angular de la base móvil con respecto al servomotor SRV02). El giroscopio es actuado a su vez por medio de dos motores de corriente directa, un motor es utilizado para desplazar el marco externo (Motor 1) y el segundo motor hace girar a velocidad constante el disco del giroscopio (Motor 2). Por tal motivo, sólo se cuenta con un motor para el control de los dos grados de libertad del giroscopio y es la razón por la que este prototipo se considera un sistema mecánico subactuado, en el sentido que tiene más grados de libertad que entradas de control.

### B. Modelo dinámico

En esta sección se muestra el modelo dinámico del giroscopio de tres grados de libertad basado en las ecuaciones de

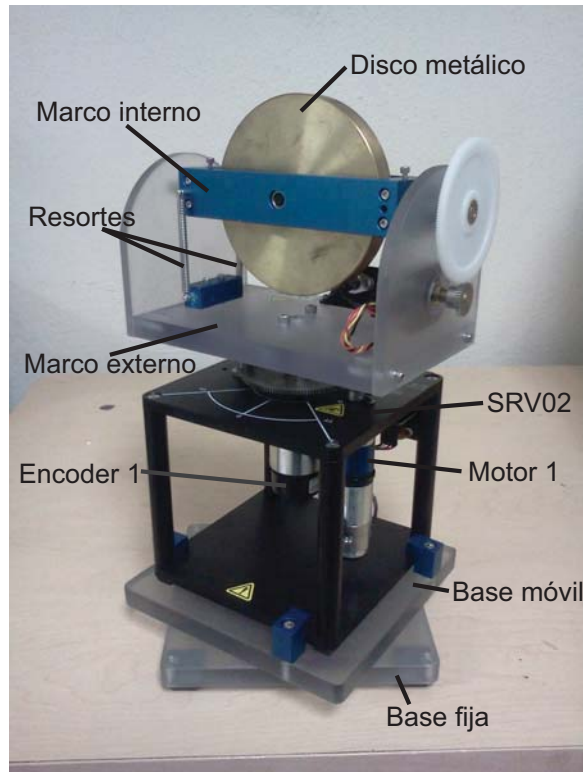


Fig. 1. Giroscopio Quanser de 3 grados de libertad

Euler-Lagrange tal y como se presentó en [1]. El estudio inicia con un análisis cinemático, donde la estrategia que se sigue es unir marcos móviles  $o_i x_i y_i z_i$  a cada eslabón de la cadena cinemática y posteriormente relacionar la orientación de cada marco móvil respecto a un marco fijo  $o_0 x_0 y_0 z_0$  localizado en el centro de masa del disco mediante la matriz de rotación  $R_i^0$  con  $i = 1, 2, 3$ .

La Figura 2 es un diagrama esquemático del giroscopio en una configuración arbitraria donde se observan cada uno de los ejes de rotación, donde  $\psi$  es el ángulo asociado al eslabón 1 medido respecto al eje  $z_0$ ;  $\gamma$  es el ángulo asociado al eslabón 2 medido respecto al eje  $y_1$  y  $\eta$  es el ángulo asociado al eslabón 3 medido respecto al eje  $x_2$ . Las siguientes matrices de rotación relacionan la orientación de cada marco móvil  $o_i x_i y_i z_i$  respecto al marco fijo  $o_0 x_0 y_0 z_0$ :

$$R_1^0 = \begin{bmatrix} C_\psi & -S_\psi & 0 \\ S_\psi & C_\psi & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

$$R_2^0 = \begin{bmatrix} C_\psi C_\gamma & -S_\psi & C_\psi S_\gamma \\ S_\psi C_\gamma & C_\psi & S_\psi S_\gamma \\ -S_\gamma & 0 & C_\gamma \end{bmatrix},$$

$$R_3^0 = \begin{bmatrix} C_\psi C_\gamma & C_\psi S_\gamma S_\eta - S_\psi C_\gamma & C_\psi S_\gamma C_\eta + S_\psi S_\eta \\ S_\psi C_\gamma & S_\psi S_\gamma S_\eta + C_\psi C_\eta & S_\psi S_\gamma C_\eta - C_\psi S_\eta \\ -S_\gamma & C_\gamma S_\eta & C_\gamma C_\eta \end{bmatrix},$$

donde  $C_j = \cos(j)$  y  $S_j = \sin(j)$  con  $j = \psi, \gamma, \eta$ .

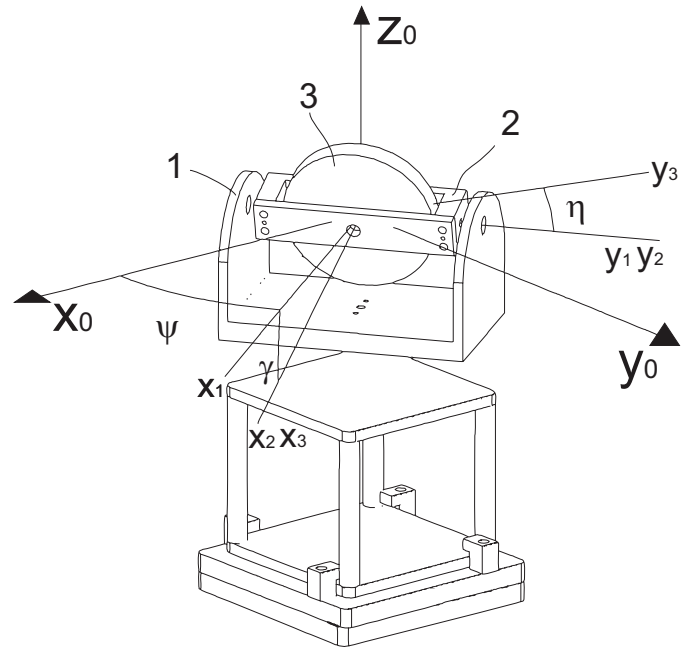


Fig. 2. Diagrama esquemático del giroscopio Quanser

La velocidad angular  $\omega_i^0$  de cada marco móvil (y por lo tanto de cada eslabón) se puede determinar a partir de las componentes de la matriz antisimétrica  $F(\omega_i^0) = -F(\omega_i^0)^T$  donde  $F(\omega_i^0)$  cumple con  $\frac{d}{dt} R_i^0 = F(\omega_i^0) R_i^0$ :

$$\omega_1^0 = \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix}, \omega_2^0 = \begin{bmatrix} -S_\psi \dot{\gamma} \\ C_\psi \dot{\gamma} \\ \dot{\psi} \end{bmatrix}, \omega_3^0 = \begin{bmatrix} -S_\psi \dot{\gamma} + C_\psi C_\gamma \dot{\eta} \\ C_\psi \dot{\gamma} + S_\psi C_\gamma \dot{\eta} \\ \dot{\psi} - S_\gamma \dot{\eta} \end{bmatrix}. \quad (1)$$

Las velocidades angulares absolutas fueron determinadas a partir de cada matriz de rotación. Ahora, se procede a calcular las energías cinética y potencial del giroscopio. Definiendo el vector de coordenadas y sus velocidades como:

$$\mathbf{q} = \begin{bmatrix} \psi \\ \gamma \\ \eta \end{bmatrix}, \dot{\mathbf{q}} = \begin{bmatrix} \dot{\psi} \\ \dot{\gamma} \\ \dot{\eta} \end{bmatrix}. \quad (2)$$

Ya que no existe traslación, la energía cinética de cada eslabón se calcula como

$$K(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \omega_i^0 T I_i^0 \omega_i^0 \quad (3)$$

donde  $I_i^0$  es la matriz de momentos de inercia del  $i$ -ésimo eslabón respecto al marco fijo  $o_0 x_0 y_0 z_0$ . Sumando las energías cinéticas de los tres eslabones y relacionando cada velocidad angular en (1) con el vector  $\dot{\mathbf{q}}$  en (2), se puede expresar a  $K(\mathbf{q}, \dot{\mathbf{q}})$  en forma compacta

$$K(\mathbf{q}, \dot{\mathbf{q}}) = \frac{1}{2} \dot{\mathbf{q}}^T M(\mathbf{q}) \dot{\mathbf{q}} \quad (4)$$

con

$$M(\mathbf{q}) = \begin{bmatrix} I_{bz} + I_z \cos^2 \gamma + I_x \sin^2 \gamma & 0 & -I_{dx} \sin \gamma \\ 0 & I_y & 0 \\ -I_{dx} \sin \gamma & 0 & I_{dx} \end{bmatrix}. \quad (5)$$

Debido al par de resortes que acoplan el marco interior con el marco exterior, la energía potencial  $U(\mathbf{q})$  del sistema es

$$U(\mathbf{q}) = \frac{1}{2} K_r (x - \gamma r)^2 + \frac{1}{2} K_r (x + \gamma r)^2 \quad (6)$$

donde  $r$  es la distancia entre el eje  $y_1$  y el resorte,  $x$  representa la deformación del resorte en metros y  $K_r$  es la constante de rigidez del resorte.

Basado en las ecuaciones de Euler-Lagrange, un modelo dinámico del giroscopio puede escribirse como:

$$M(\mathbf{q})\ddot{\mathbf{q}} + C(\mathbf{q}, \dot{\mathbf{q}})\dot{\mathbf{q}} + \mathbf{g}(\mathbf{q}) = G\mathbf{u} \quad (7)$$

con

$$C(\mathbf{q}, \dot{\mathbf{q}}) = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix}, \mathbf{g}(\mathbf{q}) = \begin{bmatrix} 0 \\ 2K_r r^2 \gamma \\ 0 \end{bmatrix}, G = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix},$$

$$\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

donde  $M(\mathbf{q})$  viene dada por (5) y la matriz  $C(\mathbf{q}, \dot{\mathbf{q}})$  se determina a partir de los símbolos de Christoffel (del primer tipo),

$$\begin{aligned} C_{11} &= (I_x - I_z) \cos \gamma \sin \gamma \dot{\gamma}, \\ C_{12} &= (I_x - I_z) \cos \gamma \sin \gamma \dot{\psi} - \frac{1}{2} I_{dx} \cos \gamma \dot{\eta}, \\ C_{13} &= -\frac{1}{2} I_{dx} \cos \gamma \dot{\gamma}, \\ C_{21} &= (I_z - I_x) \cos \gamma \sin \gamma \dot{\psi} + \frac{1}{2} I_{dx} \cos \gamma \dot{\eta}, \\ C_{22} &= 0, \\ C_{23} &= \frac{1}{2} I_{dx} \cos \gamma \dot{\psi}, \\ C_{31} &= -\frac{1}{2} I_{dx} \cos \gamma \dot{\gamma}, \\ C_{32} &= -\frac{1}{2} I_{dx} \cos \gamma \dot{\psi}, \\ C_{33} &= 0. \end{aligned}$$

Debido a que  $\text{rango}[G] = 2$  el giroscopio es un sistema mecánico subactuado en el sentido que tiene tres grados de libertad y dos actuadores. Dado que  $M(\mathbf{q})$  es invertible, el

modelo dinámico del giroscopio puede escribirse en términos del vector de estado  $[\mathbf{q} \ \dot{\mathbf{q}}]^T$  como:

$$\frac{d}{dt} \begin{bmatrix} \psi \\ \gamma \\ \eta \\ \dot{\psi} \\ \dot{\gamma} \\ \dot{\eta} \end{bmatrix} = \begin{bmatrix} \dot{\psi} \\ \dot{\gamma} \\ \dot{\eta} \\ \frac{u_1 + u_2 \sin \gamma + \cos \gamma \dot{\gamma} (I_{dx} \dot{\eta} + (I_{dx} - 2I_x + 2I_z) \dot{\psi} \sin \gamma)}{I_{bz} + I_z \cos^2 \gamma + (I_x - I_{dx}) \sin^2 \gamma} \\ -\frac{2K_r r^2 \gamma + \cos \gamma \dot{\psi} (I_{dx} \dot{\eta} + (I_z - I_x) \sin \gamma \dot{\psi})}{I_y} \\ \frac{I_{dx} \cos \gamma \dot{\gamma} a_1(\mathbf{q}, \dot{\mathbf{q}}) + 2a_2(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u})}{2I_{dx} (I_{bz} + I_z \cos^2 \gamma + (I_x - I_{dx}) \sin^2 \gamma)} \end{bmatrix}, \quad (8)$$

donde  $a_1(\mathbf{q}, \dot{\mathbf{q}})$  y  $a_2(\mathbf{q}, \dot{\mathbf{q}})$  están dadas por

$$\begin{aligned} a_1(\mathbf{q}, \dot{\mathbf{q}}) &= (2I_{bz} - I_x + 3I_z + (I_x - I_z) \cos[2\gamma]) \dot{\psi} + 2I_{dx} \dot{\eta} \sin \gamma, \\ a_2(\mathbf{q}, \dot{\mathbf{q}}, \mathbf{u}) &= u_1 I_{dx} \sin \gamma + u_2 (I_{bz} + I_z \cos^2 \gamma + I_x \sin^2 \gamma). \end{aligned}$$

1) *Modelo lineal:* Tomando en cuenta la dinámica de los motores para considerar las entradas de control en términos de voltajes en lugar de pares, se tiene que ([7]):

$$u_1 = \frac{n_{m1} k_{t1} n_{g1} k_{g1}}{R_{m1}} (V_{m1} - k_{m1} k_{g1} \dot{\psi}), \quad (9)$$

$$u_2 = \frac{n_{m2} k_{t2} n_{g2} k_{g2}}{R_{m2}} (V_{m2} - k_{m2} k_{g2} \dot{\eta}), \quad (10)$$

donde  $n_{m_i}, k_{t_i}, n_{g_i}, k_{g_i}$  and  $R_{m_i}$  son los parámetros mecánicos y eléctricos de cada uno de los motores, con  $i = 1, 2$ , cuyos valores numéricos son mostrados en la Tabla 1. Ahora, si consideramos que la velocidad  $\dot{\eta}$  del disco gira a

TABLE I  
PARÁMETROS FÍSICOS

Parámetro	Valor	Unidad
$I_{dx}$	0.001032	$kg \cdot m^2$
$I_y$	0.000544	$kg \cdot m^2$
$I_t$	0.002	$kg \cdot m^2$
$r$	0.0254	$m$
$k_{r1}$	1908.9	$N/m$
$k_{t1}$	0.00768	$N \cdot m/A$
$k_{m1}$	0.00768	$V/(rad/s)$
$R_{m1}$	2.6	$\Omega$
$k_{g1}$	70	Adimensional
$n_{g1}$	0.90	Adimensional
$n_{m1}$	0.69	Adimensional
$B_g$	0.015	$N \cdot m/(rad/s)$

una velocidad constante  $\bar{\omega}_\eta$ , se puede verificar que conjunto de equilibrios no aislados del sistema en malla cerrada es:

$$\begin{aligned} \mathcal{E} &= \{[\mathbf{q}^T \ \dot{\mathbf{q}}^T]^T \in \mathbb{R}^6 : [\psi \ \dot{\psi} \ \gamma \ \dot{\gamma} \ \eta]^T \\ &= [0 \ 0 \ 0 \ 0 \ \bar{\omega}_\eta]^T, \eta \in \mathbb{R}\} \end{aligned} \quad (11)$$

con  $V_{m1} = 0$  y  $V_{m2} = k_{m2} k_{g2} \bar{\omega}_\eta$ . Dado que el giroscopio viene diseñado para que el disco siempre gire a una velocidad  $\bar{\omega}_\eta$ , siendo  $\bar{\omega}_\eta = -457$  [rad/s] donde el signo negativo denota un giro en sentido horario respecto al eje  $x_2$ , se puede linealizar el sistema en lazo cerrado alrededor de un punto



de equilibrio  $\mathcal{E}$  y obtener un modelo lineal, el cual puede reducirse a un sistema de dos grados de libertad dado por

$$\frac{d}{dt} \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{K_1}{I_y} & 0 & 0 & \frac{-h}{I_y} \\ 0 & 0 & 0 & 1 \\ 0 & \frac{h}{I_t} & 0 & -\frac{K_2}{I_t} \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ K_3 \end{bmatrix} V_{m1}, \quad (12)$$

donde  $[x_1 \ \dot{x}_1 \ x_2 \ \dot{x}_2]^T = [\gamma \ \dot{\gamma} \ \psi \ \dot{\psi}]^T$ ,  $h = I_{dx}\bar{\omega}_\eta$ ,  $K_1 = 2k_r r^2$ ,  $K_2 = B_g + \frac{n_{m1}k_{t1}n_{g1}k_{g1}^2k_{m1}}{R_{m1}}$  y  $K_3 = \frac{n_{m1}k_{t1}n_{g1}k_{g1}}{I_t R_{m1}}$ . En nuestro caso, se ha considerado como variable de salida al desplazamiento angular  $x_2 = \psi$ , esto es,

$$y = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ \dot{x}_1 \\ x_2 \\ \dot{x}_2 \end{bmatrix} \quad (13)$$

2) *Objetivo de control*: El objetivo de control es lograr que el error del ángulo de desplazamiento  $\tilde{\psi} = \psi_d - \psi$ , del marco exterior con respecto al eje de giro de la velocidad del disco, se desvanezca, donde  $\psi_d$  es un ángulo de desplazamiento deseado, esto es,

$$\lim_{t \rightarrow \infty} \tilde{\psi}(t) = 0 \quad (14)$$

a partir de un conjunto de condiciones iniciales tan grande como sea posible.

### III. DISEÑO DE UN REGULADOR CUADRÁTICO LINEAL (LQR)

A partir de la forma canónica de un sistema lineal dado por

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}v_1 \quad (15)$$

donde  $v_1 = V_{m1}$ , se define su salida como:

$$\psi = C_z \mathbf{x}, \quad (16)$$

con

$$C_z = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}.$$

Ahora, tomando en cuenta la definición el error de desplazamiento angular  $\tilde{\psi}$  para el diseño del regulador, al derivar  $\tilde{\psi}$  con respecto al tiempo se obtiene:

$$\dot{\tilde{\psi}} = -\dot{\psi} \quad (17)$$

ya que  $\psi_d$  es constante. Sustituyendo (16) en (17) se consigue:

$$\dot{\tilde{\psi}} = -C_z \dot{\mathbf{x}}. \quad (18)$$

Introduciendo un cambio de variable  $\mathbf{z} = \dot{\mathbf{x}}$ , (18) se reescribe como:

$$\dot{\tilde{\psi}} = -C_z \mathbf{z}. \quad (19)$$

Derivando  $\mathbf{z}$  con respecto al tiempo se obtiene:

$$\dot{\mathbf{z}} = \ddot{\mathbf{x}} = \mathbf{A}\dot{\mathbf{x}} + \mathbf{B}\dot{v}_1, \quad (20)$$

ya que  $A$  y  $B$  son matrices con coeficientes constantes. Una vez más, introducimos un cambio de variable  $w = \dot{v}_1$ , y al reescribir (20) en términos de  $\mathbf{z}$  y  $w$  se obtiene:

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}w. \quad (21)$$

Ahora, definimos el vector  $\bar{\mathbf{x}} = [\mathbf{z}^T \ \tilde{\psi}]^T \in \mathbb{R}^4$  para obtener la siguiente ecuación de estado:

$$\frac{d}{dt} \begin{bmatrix} \mathbf{z} \\ \tilde{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{0}_{4 \times 1} \\ -C_z & 0 \end{bmatrix} \bar{\mathbf{x}} + \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix} w, \quad (22)$$

$$\dot{\bar{\mathbf{x}}} = \bar{\mathbf{A}}\bar{\mathbf{x}} + \bar{\mathbf{B}}w.$$

Usando Matlab se obtiene la matriz de controlabilidad del sistema (22) la cual es de rango pleno. Lo anterior significa que se puede diseñar una ley de control dada por

$$w = -[\mathbf{K}_s \quad k_i] \begin{bmatrix} \mathbf{z} \\ \tilde{\psi} \end{bmatrix}$$

o en forma compacta,

$$w = -\bar{\mathbf{K}}\bar{\mathbf{x}}, \quad (23)$$

tal que al sustituirla en (22) se obtiene el sistema en lazo cerrado

$$\frac{d}{dt} \begin{bmatrix} \mathbf{z} \\ \tilde{\psi} \end{bmatrix} = \begin{bmatrix} \mathbf{A} - \mathbf{B}\mathbf{K}_s & -\mathbf{B}k_i \\ -C_z & 0 \end{bmatrix} \begin{bmatrix} \mathbf{z} \\ \tilde{\psi} \end{bmatrix}$$

o bien,

$$\dot{\bar{\mathbf{x}}} = (\bar{\mathbf{A}} - \mathbf{B}\bar{\mathbf{K}})\bar{\mathbf{x}}, \quad (24)$$

$$\dot{\tilde{\psi}} = -C_z \mathbf{z} \quad (25)$$

El diseño del LQR se lleva a cabo mediante el comando *lqr* de Matlab, donde se obtiene

$$\mathbf{Q} = \begin{bmatrix} 0.1801 & 0 & 0 & 0 & 0 \\ 0 & 0.0028 & 0 & 0 & 0 \\ 0 & 0 & 0.1821 & 0 & 0 \\ 0 & 0 & 0 & 0.0028 & 0 \\ 0 & 0 & 0 & 0 & 22.19 \end{bmatrix},$$

$$R = 1,$$

y

$$\bar{\mathbf{K}} = \begin{bmatrix} 10.8694 & -0.0033 & 5.2428 & 0.1243 & \vdots & -4.7739 \end{bmatrix},$$

$$= \begin{bmatrix} \mathbf{K}_s & \vdots & k_i \end{bmatrix}.$$

Los polos en lazo cerrado son:  $p_1 = -9.13 + 303.69i$ ,  $p_2 = -9.13 - 303.69i$ ,  $p_3 = -0.97 + 1.55i$ ,  $p_4 = -0.97 - 1.55i$  y  $p_5 = -1.93$ . Ahora, recordando que el LQR se diseñó con  $w = \dot{v}_1$ , se procede a integrar (23) para tener la ley de control en términos de  $v_1$ , esto es,

$$v_1 = -\mathbf{K}_s \mathbf{x} - k_i \int_0^t \tilde{\psi}(\tau) d\tau. \quad (26)$$

Con la medición completa el vector de estado  $\mathbf{x}$  es posible implementar la ley de control (26). Sin embargo, como se mencionó anteriormente, esto no es posible debido a que el giroscopio sólo cuenta con sensores de posición. Por tanto, reescribimos (26) como

$$v_1 = -\mathbf{K}_s \hat{\mathbf{x}} - k_i \int_0^t \tilde{\psi}(\tau) d\tau, \quad (27)$$

donde  $\hat{x}$  es el vector de estados estimados. A continuación se presenta el análisis de estabilidad del sistema (15) con la ley de control (27) y un observador de estados de orden completo

#### A. Diseño del observador

Se propone un observador de estados de orden completo de la forma

$$\dot{\hat{x}} = A\hat{x} + Bv_1 + L\tilde{y}, \quad (28)$$

para calcular las velocidades del giroscopio a partir de los dos sensores de posición, donde  $\hat{x}$  es el vector de estado estimado y  $\tilde{y} := y - \hat{y}$ . Así,

$$y = C_y x$$

y

$$\hat{y} = C_y \hat{x}$$

con

$$C_y = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

Definiendo el error de estimación como  $\tilde{x} := x - \hat{x}$ , entonces su dinámica está dada por

$$\dot{\tilde{x}} = (A - LC_y)\tilde{x}, \quad (29)$$

el cual es un sistema autónomo. Siempre que (29) tenga polos con parte real negativa, el error de seguimiento será asintóticamente estable. Sobre la ubicación de los polos del observador de estados, es recomendable colocarlos entre dos y cinco constantes de tiempo más alejados de los polos del sistema en lazo cerrado. Se seleccionan los siguientes polos para el observador de estados de orden completo:  $o_1 = -30 + 10i$ ,  $o_2 = -30 - 10i$ ,  $o_3 = -25 + 5i$  y  $o_4 = -25 - 5i$ . Mediante el comando *place* de Matlab se calcula la ganancia del observador, obteniendo:

$$L = \begin{bmatrix} 52.43 & 884.62 \\ -91362 & 34781 \\ -103.92 & 38.99 \\ -3526.4 & -92002 \end{bmatrix}. \quad (30)$$

Finalmente, con la matriz de ganancia del observador, se puede implementar la ley de control (27).

#### B. Estabilidad

Se procede a probar la estabilidad del sistema en lazo cerrado con la ley de control (27) y el observador de estados (28). Sustituyendo la ley de control (27) en (15) se obtiene:

$$\dot{x} = Ax - BK_s \hat{x} - Bk_i \int_0^t \tilde{\psi}(\tau) d\tau. \quad (31)$$

A partir de la definición  $\tilde{x} := x - \hat{x}$  se despeja  $\hat{x}$  y se sustituye en (31) para obtener:

$$\dot{x} = (A - BK_s)x + BK_s \tilde{x} - Bk_i \int_0^t \tilde{\psi}(\tau) d\tau. \quad (32)$$

Examinando (32) observamos que la dinámica del sistema en lazo cerrado depende de  $x$ ,  $\tilde{x}$  y  $\int_0^t \tilde{\psi}(\tau) d\tau$ . Debido a que en

(32) no aparece  $\tilde{\psi}$  sino su integral, se deriva esta ecuación respecto al tiempo, obteniendo:

$$\dot{\tilde{x}} = (A - BK_s)\tilde{x} + BK_s \dot{\tilde{x}} - Bk_i \dot{\tilde{\psi}}. \quad (33)$$

el cual es un sistema dinámico que depende de  $\dot{\tilde{\psi}}$ . Tomando en cuenta  $z = \dot{\tilde{x}}$  y  $\tilde{z} = \dot{\tilde{x}}$  la ecuación (33) se reescribe como:

$$\dot{\tilde{z}} = (A - BK_s)\tilde{z} + BK_s \dot{\tilde{z}} - Bk_i \dot{\tilde{\psi}}. \quad (34)$$

Al derivar (29) respecto al tiempo y sustituyendo  $\tilde{z} = \dot{\tilde{x}}$  se obtiene:

$$\dot{\tilde{z}} = (A - LC_y)\tilde{z}. \quad (35)$$

Con (34), (35) y (19) se escribe la dinámica de lazo cerrado en términos del vector de estado  $\tilde{z} = [z^T \ \tilde{z}^T \ \tilde{\psi}]^T$ :

$$\underbrace{\frac{d}{dt} \begin{bmatrix} z \\ \tilde{z} \\ \tilde{\psi} \end{bmatrix}}_{\dot{\tilde{z}}} = \underbrace{\begin{bmatrix} A - BK_s & BK_s & -Bk_i \\ \mathbf{0}_{4 \times 4} & A - LC_y & \mathbf{0}_{4 \times 1} \\ -C_z & \mathbf{0}_{1 \times 4} & 0 \end{bmatrix}}_{\hat{A}} \underbrace{\begin{bmatrix} z \\ \tilde{z} \\ \tilde{\psi} \end{bmatrix}}_{\tilde{z}}, \quad (36)$$

el cual en notación compacta queda de la siguiente forma

$$\dot{\tilde{z}} = \hat{A}\tilde{z}, \quad (37)$$

Puede verificarse que (37) es asintóticamente estable, tal que el objetivo de control (14) se cumple.

#### C. Resultados experimentales

En la Figura 3 se muestra un diagrama del sistema en lazo cerrado. El experimento se realizó con Matlab/Simulink usando la librería QUARC de Quanser. El experimento inicia

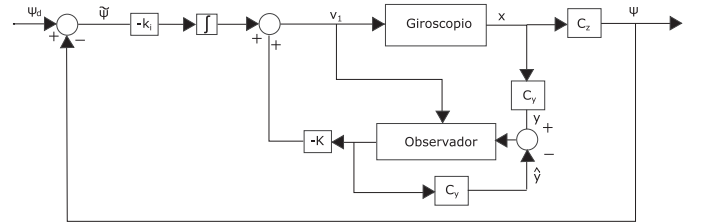


Fig. 3. Diagrama de bloques control LQR.

con una condición inicial del giroscopio diferente al del observador de estado. Los resultados experimentales se muestran en la Figura 4. Puede observarse que se cumple el objetivo de control (14), esto es, logramos llevar el giroscopio a un desplazamiento angular deseado sin exceder el voltaje nominal de  $\pm 6V$  del motor.

#### IV. CONCLUSIONES

Se ha presentado en detalle el diseño de un controlador LQR para el control de un giroscopio de tres grados de libertad de la marca Quanser, basado en el modelo dinámico obtenido en [1]. Para validar nuestro sistema de control se mostraron resultados experimentales donde logramos llevar el marco exterior del giroscopio a un desplazamiento angular deseado.

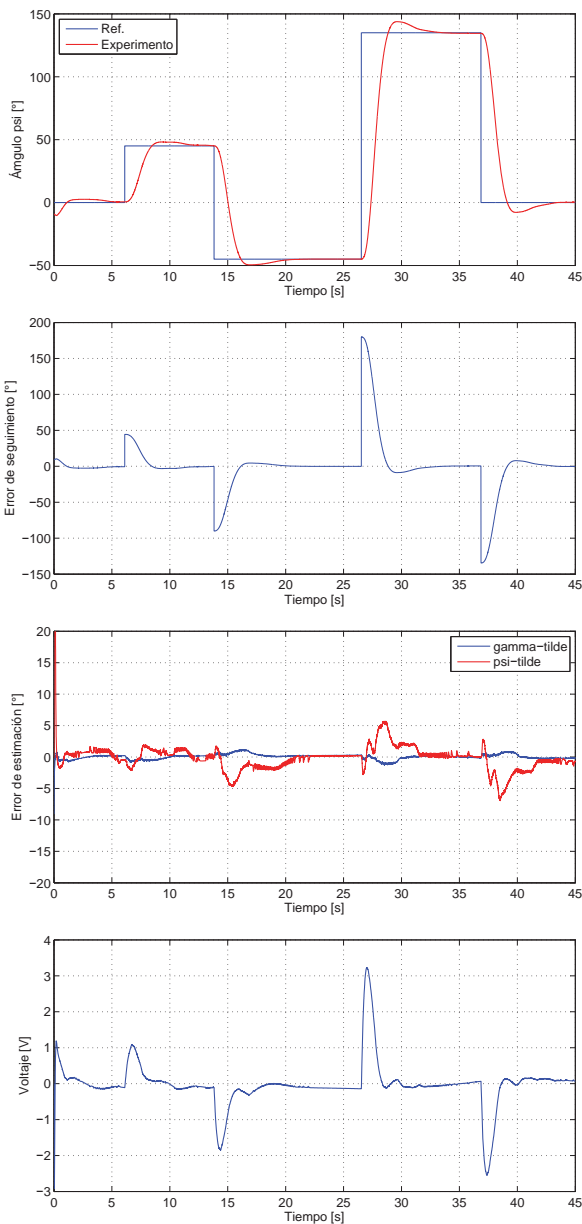


Fig. 4. Resultados experimentales.

#### AGRADECIMIENTOS

Este trabajo fue parcialmente apoyado por el Tecnológico Nacional de México (Contratos TecNM 5345.14-P y 5574.15-P), y por el Consejo Nacional de Ciencia y Tecnología (Contratos CONACyT 166636 y 134534).

#### REFERENCIAS

- [1] J. Moyrón, J. A. Sandoval. *Control de un giroscopio basado en un compensador en adelanto*, XVII Congreso Mexicano de Robótica, pp. 145-150, Noviembre 2015, ISBN:978-607-97128-0-8.
- [2] G. Cordero, A. Dzul, V. Santibáñez, E. J. Ollervides. *2014 Sintonización de un controlador PID-PID en cascada aplicado a un giroscopio de 2 g.d.l.* XVI Congreso Latinoamericano de Control Automático, CLCA 2014.

- [3] A.S. Chandak, A.J. Patil. *2013 Robust LQR Control Design of Gyroscope* International Journal of Advanced Computer Research, Volume-3, Number-1, Issue-9, March-2013.
- [4] J.L. Mora, E. Ruíz. *2008 Robust stability of a gyroscope using linear matrix inequalities*. Conciencia Tecnológica, núm. 36, julio-diciembre, pp. 29-34.
- [5] Quanser. *2012 USER MANUAL Gyro/Stable Platform Experiment*. v1.0
- [6] Quanser. *2012 USER MANUAL SRV02 Rotary Servo Base Unit*. v1.0
- [7] K. Ogata. *2010 Modern Control Engineering*. Fifth Edition, Prentice Hall.

# Implementación de algoritmos de control en tiempo real en la plataforma robótica Segway RMP-100

Israel Soto

Instituto de Ingeniería y Tecnología  
 Universidad Autónoma de Ciudad Juárez  
 Ciudad Juárez 32310, Mexico  
 Email: angel.soto@uacj.mx

Eusebio Bugarin

Tecnológico Nacional de México  
 Instituto Tecnológico de Ensenada  
 Ensenada 22750, Mexico.

Alfredo Delgado  
y Ricardo Campa

Tecnológico Nacional de México  
 Instituto Tecnológico de la Laguna  
 Torreón 27000, Mexico.

**Resumen**—En este trabajo se aborda el modelado y control de una plataforma robótica móvil de Segway Inc. El modelo particular bajo estudio (RMP-100) es un péndulo invertido sobre ruedas que tiene sólo dos ruedas coaxiales actuadas de forma independiente para realizar sus dos tareas principales: mantener el péndulo invertido en su posición vertical (balanceo), y mover el robot a una posición y orientación específicas sobre el piso (control de postura). Se inicia revisando algunos temas sobre el modelado de la dinámica de sistemas restringidos y se obtiene el modelo dinámico del prototipo utilizando la formulación de Lagrange; posteriormente, se describe la implementación de algunos algoritmos de control tanto para el balanceo como para el movimiento de la plataforma, y al final se incluyen resultados experimentales.

## I. INTRODUCCIÓN

Los sistemas tipo péndulo invertido sobre ruedas (PIR) han sido estudiados en muchos trabajos recientes de la comunidad científica, aunque los primeros trabajos relacionados con el tema datan de los principios de 1940 [1]. El interés actual en el estudio de este tipo de robot móvil se debe principalmente a la forma de su estructura mecánica (con sólo dos ruedas colocadas sobre el mismo eje, y cada una actuada por un motor independiente) y su amplio rango de aplicaciones tales como la exploración, la vigilancia o el transporte. Además, desde el punto de vista de la teoría de control, los PIR poseen retos interesantes. Por ejemplo, un PIR no sólo es un sistema no holonómico, sino también un sistema subactuado. Y si el control de balanceo se desprecia, un PIR puede ser reducido a un robot móvil diferencial (también conocido como robot móvil tipo unicycle).

De acuerdo con la literatura, para el control de un PIR generalmente se emplea un control por realimentación de estados lineal (que hace uso del modelo dinámico linealizado) [1]–[4], o bien un control por modos deslizantes [5]. Estudios más recientes han analizado este sistema robótico considerando su modelo dinámico no lineal con el fin de diseñar un controlador no lineal (ver [6] y [7]).

Hoy en día, existen diferentes robots tipo PIR para uso académico, como es el caso de Joe [3], Quasimoro [4], y los que produce y vende la compañía Segway desde el año 2000. El prototipo estudiado en este artículo es precisamente un robot móvil tipo PIR (modelo RMP-100) de Segway.

El objetivo de este artículo es doble. Primero, después de recordar algunos aspectos sobre modelado dinámico de sistemas con restricciones, se presenta un procedimiento para calcular el modelo dinámico del robot bajo estudio. Luego, se explica cómo implementar en tal plataforma dos controladores en tiempo real: un controlador de seguimiento de postura y un controlador de balanceo para el péndulo; y se llevan a cabo algunos experimentos para verificar la viabilidad de la metodología propuesta.

## II. MODELADO DE SISTEMAS RESTRINGIDOS

Se le llama sistema restringido a un sistema mecánico el cual puede ser modelado utilizando  $m$  coordenadas generalizadas  $\rho_1, \rho_2, \dots, \rho_m$ , pero esas coordenadas (y/o sus velocidades) no son independientes, en el sentido de que se pueden encontrar algunas restricciones holonómicas o no holonómicas entre ellas.

Sea  $\rho$  el vector de coordenadas generalizadas, i.e.,  $\rho = [\rho_1 \ \rho_2 \ \dots \ \rho_m]^T$ . Si el sistema tiene sólo  $n$  grados de libertad (g.d.l.), con  $n < m$ , entonces existen  $p = m - n$  restricciones holonómicas de la forma:

$$\mu_i(\rho) = 0, i = 1, 2, \dots, p; \quad (1)$$

y se debe hacer notar que cualquier restricción holonómica implica también una restricción en las velocidades generalizadas, dadas por el vector  $\dot{\rho}$ , de modo que  $\dot{\mu}_i(\rho, \dot{\rho}) = 0$ . Este tipo de restricciones es común en robots manipuladores paralelos.

Una restricción no holonómica, por otra parte, es una restricción que existe en las velocidades generalizadas, pero no en las coordenadas generalizadas. Esto es, existen algunas relaciones de la forma

$$\nu_i(\rho, \dot{\rho}) = 0, i = 1, 2, \dots, p; \quad (2)$$

pero estas relaciones no pueden ser integradas para obtener las restricciones de la forma (1). En otras palabras, una restricción no holonómica no reduce el espacio de configuración del sistema, sino sólo el espacio de velocidades instantáneas. Este tipo de restricciones usualmente aparecen en robots móviles con ruedas, aunque en este caso particular las restricciones (2) pueden ser agrupadas de manera que quede

$$A(\rho)\dot{\rho} = 0 \quad (3)$$

donde  $A(\rho) \in \mathbb{R}^{p \times m}$  es la matriz de restricciones no holonómicas pfaffianas.

El modelo dinámico de un mecanismo con restricciones pfaffianas pueden ser representado por el siguiente conjunto de ecuaciones algebro-diferenciales:

$$\frac{d}{dt} \left\{ \frac{\partial \mathcal{L}(\rho, \dot{\rho})}{\partial \dot{\rho}} \right\} - \frac{\partial \mathcal{L}(\rho, \dot{\rho})}{\partial \rho} = \tau_\rho + A(\rho)^T \lambda \quad (4)$$

donde el lagrangiano  $\mathcal{L}(\rho, \dot{\rho})$  se define como

$$\mathcal{L}(\rho, \dot{\rho}) = K(\rho, \dot{\rho}) - U(\rho) \quad (5)$$

siendo  $K(\rho, \dot{\rho})$  la energía cinética total y  $U(\rho)$  la energía potencial total del mecanismo;  $\tau_\rho \in \mathbb{R}^m$  es el vector de fuerzas generalizadas, y  $\lambda \in \mathbb{R}^p$  es el vector de multiplicadores de Lagrange, los cuales aseguran el cumplimiento de las restricciones dadas por (3). La ecuación (4) puede ser reescrita como

$$M(\rho)\ddot{\rho} + C(\rho, \dot{\rho})\dot{\rho} + g(\rho) = \tau_\rho + A(\rho)^T \lambda \quad (6)$$

donde  $M(\rho) \in \mathbb{R}^{m \times m}$  es una matriz simétrica y definida positiva conocida como matriz de inercias,  $C(\rho, \dot{\rho}) \in \mathbb{R}^{m \times m}$  es la matriz de términos generados por fuerzas centrífugas y de Coriolis,  $g(\rho) \in \mathbb{R}^m$  es el vector de fuerzas debidas a la gravedad. Además, como se explica en [8], se satisfacen las siguientes propiedades para  $M(\rho)$  y  $g(\rho)$ :

$$K(\rho, \dot{\rho}) = \frac{1}{2} \dot{\rho}^T M(\rho) \dot{\rho} \quad (7)$$

$$g(\rho) = \frac{\partial U(\rho)}{\partial \rho} \quad (8)$$

y dada la matriz  $M(\rho)$ , es posible calcular la matriz  $C(\rho, \dot{\rho})$  empleando los símbolos de Christoffel (ver [8]).

Con el fin de proyectar la dinámica (6) al espacio tangente de la variedad de restricción (3) se necesita encontrar una matriz  $R(\rho) \in \mathbb{R}^{m \times n}$  cuyo espacio columna pertenezca al espacio nulo de  $A(\rho)$ , i.e.,  $A(\rho)R(\rho) = \mathbf{0}$ . Entonces todas las velocidades dependientes factibles  $\dot{\rho}$  del sistema restringido pertenecen al espacio generado por las columnas de  $R(\rho)$ , el cual está parametrizado por un vector de velocidades independientes  $\dot{q} \in \mathbb{R}^n$ , lo que lleva a la siguiente expresión

$$\dot{\rho} = R(\rho)\dot{q}. \quad (9)$$

Nótese que para un sistema particular la matriz  $R(\rho)$  depende de la selección de los vectores de coordenadas generalizadas  $\rho$  y de coordenadas independientes  $q$ .

Utilizando tal matriz  $R(\rho)$ , es posible reducir el sistema (6) a un nuevo sistema dado por

$$M_r(\rho)\ddot{q} + C_r(\rho, \dot{q})\dot{q} + g_r(\rho) = \tau_r, \quad (10)$$

donde

$$\begin{aligned} M_r(\rho) &= R(\rho)^T M(\rho) R(\rho) \\ C_r(\rho, \dot{q}) &= R(\rho)^T C(\rho, \dot{\rho}) R(\rho) + R(\rho)^T M(\rho) \dot{R}(\rho, \dot{q}) \\ g_r(\rho) &= R(\rho)^T g(\rho) \\ \tau_r &= R(\rho)^T \tau_\rho \end{aligned}$$

y nótese que los estados que definen la dinámica del sistema reducido están dados por  $q$ ,  $\dot{q}$  y  $\rho$ , y pueden ser obtenidos integrando las correspondientes ecuaciones de estado.

### III. LA PLATAFORMA SEGWAY RMP-100

Las plataformas robóticas móviles RMPs (de Segway) son plataformas de propulsión eléctrica para aplicaciones robóticas terrestres; entre éstas se incluye la RMP-100 (ver Fig. 1), que es estudiada en este trabajo.



Figura 1: La plataforma Segway RMP-100

#### III-A. Descripción del sistema

La plataforma RMP-100 puede ser utilizada para transportar material o equipo con un peso de hasta 45 kg, y está lista para ser integrada con un procesador de control externo utilizando protocolos de comunicación como USB (universal serial bus) o CAN (controller area network) bus. De esta manera, el robot acepta velocidades lineal y angular como entradas de referencia, con el fin de establecer su movimiento deseado, y proporciona lecturas de alguno datos, tales como la distancia recorrida, su orientación y la inclinación del péndulo (así como las derivadas temporales de estos datos) con un periodo de muestreo de 0.01s de acuerdo con el fabricante. Además, esta plataforma robótica tiene la alternativa de utilizar un controlador de autobalanceo interno, el cual mantiene la estructura del péndulo en una posición vertical.

#### III-B. Modelo Cinemático

Tomando en cuenta la Fig. 2 y considerando como vector de coordenadas generalizadas  $\rho = [x \ y \ \theta \ \alpha \ \phi_l \ \phi_r]^T \in \mathbb{R}^6$ , donde  $x$ ,  $y$  y  $\theta$  indican la posición y orientación, de la plataforma sobre el plano,  $\alpha$  denota la inclinación del péndulo, mientras que  $\phi_l$  y  $\phi_r$  son el desplazamiento angular de las ruedas izquierda y derecha, respectivamente; entonces la cinemática del mecanismo puede ser reescrita como

$$\dot{\rho} = \begin{bmatrix} \frac{rC_\theta}{2} & \frac{rS_\theta}{2} & \frac{r}{2L} & 0 & 0 & 1 \\ \frac{rC_\theta}{2} & \frac{rS_\theta}{2} & -\frac{r}{2L} & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T \begin{bmatrix} \dot{\phi}_l \\ \dot{\phi}_r \\ \dot{\alpha} \end{bmatrix}. \quad (11)$$

donde  $S_\theta$  y  $C_\theta$  denotan  $\sin(\theta)$  y  $\cos(\theta)$ , respectivamente. Ahora sean  $v$  y  $\omega$  la magnitud de las velocidades lineal y

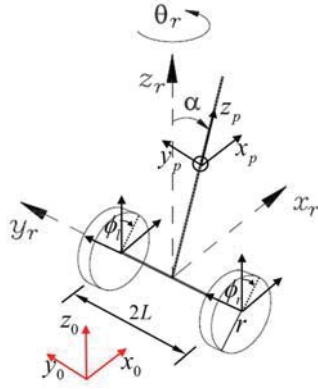


Figura 2: Diagrama cinemático

angular del robot respectivamente, entonces se puede verificar que las siguientes relaciones se dan

$$\begin{bmatrix} \dot{\phi}_l \\ \dot{\phi}_r \end{bmatrix} = \begin{bmatrix} \frac{1}{r} & -\frac{L}{r} \\ \frac{1}{r} & \frac{L}{r} \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}$$

y la ecuación (11) puede ser reescrita como

$$\dot{\rho} = \begin{bmatrix} C_\theta & S_\theta & 0 & 0 & \frac{1}{r} & \frac{1}{r} \\ 0 & 0 & 1 & 0 & -\frac{L}{r} & \frac{L}{r} \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T \begin{bmatrix} v \\ \omega \\ \dot{\alpha} \end{bmatrix}$$

### III-C. Modelo dinámico

Primeramente considérese obtener la energía cinética de cada parte del PIR para obtener la energía cinética total, dada por

$$K(\rho, \dot{\rho}) = K_b + K_p + K_{wr} + K_{wl},$$

donde  $K_b$ ,  $K_p$ ,  $K_{wr}$  y  $K_{wl}$  representan la energía cinética de la base, del péndulo, de la rueda derecha y de la rueda izquierda, respectivamente.

Se puede demostrar que la energía cinética de cada parte está dada por

$$\begin{aligned} K_b &= \frac{1}{2}m_b(\dot{x}^2 + \dot{y}^2) + \frac{1}{2}I_{bz}\dot{\theta}^2 \\ K_p &= \frac{1}{2}m_p(l_p^2\dot{\alpha}^2 + \dot{x}^2 + \dot{y}^2 - l_p^2C_\alpha^2\dot{\theta}^2 + 2l_pC_\alpha C_\theta\dot{\alpha}\dot{x} \\ &\quad + 2l_pC_\alpha S_\theta\dot{\alpha}\dot{y} + 2l_pS_\alpha C_\theta\dot{\theta}\dot{y} - 2l_pS_\alpha S_\theta\dot{\theta}\dot{x} + l_p^2\dot{\theta}^2) \\ &\quad + \frac{1}{2}(I_{py}\dot{\alpha}^2 + I_{pz}\dot{\theta}^2 + I_{px}S_\alpha^2\dot{\theta}^2 - I_{pz}S_\alpha^2\dot{\theta}^2). \\ K_{wl} &= \frac{1}{2}m_l(\dot{x}^2 + \dot{y}^2 - 2L\dot{\theta}(C_\theta\dot{x} + S_\theta\dot{y}) + L^2\dot{\theta}^2) \\ &\quad + \frac{1}{2}(I_{ly}\dot{\phi}_l^2 + I_{lz}\dot{\theta}^2 + I_{lx}S_{\phi_l}^2\dot{\theta}^2 - I_{lz}S_{\phi_l}^2\dot{\theta}^2). \\ K_{wr} &= \frac{1}{2}m_r(\dot{x}^2 + \dot{y}^2 + 2L\dot{\theta}(C_\theta\dot{x} + S_\theta\dot{y}) + L^2\dot{\theta}^2) \\ &\quad + \frac{1}{2}(I_{ry}\dot{\phi}_r^2 + I_{rz}\dot{\theta}^2 + I_{rx}S_{\phi_r}^2\dot{\theta}^2 - I_{rz}S_{\phi_r}^2\dot{\theta}^2). \end{aligned}$$

Por otro lado, considérese utilizar la altura del c.d.m. de la base como referencia con el fin de calcular la energía potencial,

y suponiendo que la altura del centro de masa (c.d.m.) de las ruedas también tienen esa altura, entonces el péndulo es el único elemento que contribuye a la energía potencial, dada por

$$U(\rho) = mgl_pC_\alpha$$

Ahora, es posible calcular el lagrangiano (5) y luego aplicar (4). Como resultado se obtiene el modelo (6), donde la matriz  $M(\rho)$  está dada por

$$M(\rho) = \begin{bmatrix} m_{11} & 0 & m_{13} & m_{14} & 0 & 0 \\ 0 & m_{22} & m_{23} & m_{24} & 0 & 0 \\ m_{13} & m_{23} & m_{33} & 0 & 0 & 0 \\ m_{14} & m_{24} & 0 & m_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & m_{55} & 0 \\ 0 & 0 & 0 & 0 & 0 & m_{66} \end{bmatrix}$$

con los elementos diferentes de cero dados por

$$\begin{aligned} m_{11} &= m_b + m_p + m_l + m_r \\ m_{13} &= -m_p l_p S_\alpha S_\theta - m_l L C_\theta + m_r L C_\theta \\ m_{14} &= m_p l_p C_\alpha C_\theta \\ m_{22} &= m_b + m_p + m_l + m_r \\ m_{23} &= m_p l_p S_\alpha C_\theta - m_l L S_\theta \\ m_{24} &= m_p l_p C_\alpha S_\theta \\ m_{33} &= I_{bz} + m_p(l_p^2 - l_p^2 C_\alpha^2) + I_{pz} + (I_{px} - I_{pz})S_\alpha^2 + m_l L^2 \\ &\quad + I_{lz} + (I_{lx} - I_{lz})S_{\phi_l}^2 + m_r L^2 + I_{rz} + (I_{rx} - I_{rz})S_{\phi_r}^2 \\ m_{44} &= m_p l_p^2 + I_{py} \\ m_{55} &= I_{ly} \\ m_{66} &= I_{ry}. \end{aligned}$$

La matriz de Coriolis  $C(\rho, \dot{\rho})$ , puede ser obtenida utilizando los símbolos de Christoffel [8].

La matriz de restricciones  $A(\rho)$  en el caso de este robot está dada por

$$A(\rho) = \begin{bmatrix} -S_\theta & C_\theta & 0 & 0 & 0 & 0 \\ C_\theta & S_\theta & L & 0 & 0 & -r \\ C_\theta & S_\theta & -L & 0 & -r & 0 \end{bmatrix}.$$

La primera fila de  $A(\rho)$  corresponde a la restricción de no desplazamiento lateral, mientras que las filas dos y tres corresponden a la restricción del desplazamiento para cada rueda.

Con el fin de eliminar los multiplicadores de Lagrange se puede utilizar la matriz dada por

$$R(\rho) = \begin{bmatrix} C_\theta & S_\theta & 0 & 0 & \frac{1}{r} & \frac{1}{r} \\ 0 & 0 & 1 & 0 & -\frac{L}{r} & \frac{L}{r} \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}^T \quad (12)$$

la cual se obtiene calculando el espacio nulo de  $A(\rho)$ , de manera que  $A(\rho)R(\rho) = \mathbf{0}$ .

Utilizando esta matriz  $R(\rho)$ , es posible obtener el sistema reducido (10), donde

$$M_r(\rho) = \begin{bmatrix} m_{r11} & 0 & m_{r13} \\ 0 & m_{r22} & 0 \\ m_{r13} & 0 & m_{r33} \end{bmatrix}$$

con

$$\begin{aligned}
 m_{r11} &= \frac{1}{r^2}(I_{ly} + I_{ry}) + (m_b + m_l + m_p + m_r) \\
 m_{r12} &= 0 \\
 m_{r13} &= m_p l_p C_\alpha \\
 m_{r22} &= I_{bz} + I_{lz} + I_{pz} + I_{rz} + (m_l + m_r) L^2 \\
 &\quad + (I_{px} - I_{pz}) S_\alpha^2 + (I_{ly} + I_{ry}) \frac{L^2}{r^2} + m_p l_p^2 S_\alpha^2 \\
 m_{r23} &= 0 \\
 m_{r33} &= m_p l_p^2 + I_{py}.
 \end{aligned}$$

La matriz de Coriolis reducida está dada por

$$C_r(\boldsymbol{\rho}, \dot{\boldsymbol{\rho}}) = \begin{bmatrix} 0 & 0 & -c_2 \dot{\alpha} \\ 0 & c_1 \dot{\alpha} & c_1 \dot{\theta} \\ 0 & -c_1 \dot{\theta} & 0 \end{bmatrix}$$

con  $c_1 = (m_p l_p^2 + (I_{px} - I_{pz})) C_\alpha S_\alpha$  y  $c_2 = m_p l_p S_\alpha$ .

Y el vector de gravedad está definido como

$$g_r(\boldsymbol{\rho}) = [0 \quad 0 \quad -gm_p l_p S_\alpha]^T$$

#### IV. CONTROL EN TIEMPO REAL

Se desarrolló un sistema computacional para control en tiempo real de la plataforma RMP-100 utilizando Windows XP y una extensión de tiempo real de IntervalZero; esta es una herramienta de software muy completa, la cual trabaja junto con Windows proporcionándole algunas características de sistema operativo de tiempo real (RTOS); este sistema se ejecuta en una PC, y emplea comunicación CAN con un periodo de muestreo de 0.01 [s].

Las entradas al robot son las magnitudes de las velocidades deseadas lineal ( $v_d$ ) y angular ( $\omega_d$ ). Este sistema robótico ofrece lecturas de la distancia lineal recorrida  $x_r$ , orientación  $\theta$  e inclinación  $\alpha$  (y sus correspondientes derivadas temporales  $\dot{x}_r$ ,  $\dot{\theta}$  y  $\dot{\alpha}$ ); esto es, todo el vector de estados puede ser medido del robot.

En este trabajo se consideraron dos controladores independientemente: uno para seguimiento considerando el sistema como un robot diferencial (es decir, en modo de autobalanceo) y otro para el balanceo, sin considerar movimiento de la postura del robot.

##### IV-A. Controlador de seguimiento

Para el control de seguimiento se decidió utilizar el controlador propuesto en [9], el cual está dado por

$$\begin{aligned}
 v_d &= \xi \\
 \omega_d &= \frac{u_2 \cos(\theta) - u_1 \sin(\theta)}{\xi} \\
 \dot{\xi} &= u_1 \cos(\theta) + u_2 \sin(\theta)
 \end{aligned}$$

donde las señales  $u_1$  y  $u_2$  son calculadas como

$$\begin{aligned}
 u_1 &= \ddot{x}_d + k_{p1} \tilde{x} + k_{v1} \dot{\tilde{x}} \\
 u_2 &= \ddot{y}_d + k_{p2} \tilde{y} + k_{v2} \dot{\tilde{y}}
 \end{aligned}$$

con  $\tilde{x} = x_d - x$ ,  $\dot{\tilde{x}} = \dot{x}_d - \dot{x}$ ,  $\tilde{y} = y_d - y$ ,  $\dot{\tilde{y}} = \dot{y}_d - \dot{y}$  y las ganancias se seleccionaron de manera que  $k_{p_i} > 0$  y  $k_{v_i} > 0$ , para  $i = 1, 2$ .

##### IV-B. Controlador de balanceo

Para el control de balanceo se decidió utilizar un controlador LQR, cuya salida es el torque requerido para mantener el péndulo en su posición vertical hacia arriba. Con el fin de diseñar el controlador LQR se requiere linealizar el modelo dinámico del robot.

Considérese  $\mathbf{z} = [\mathbf{q} \quad \dot{\mathbf{q}}]^T \in \mathbb{R}^6$  como el vector de variables de estado, donde  $\mathbf{q} = [x_r \quad \theta \quad \alpha]^T \in \mathbb{R}^3$  y  $\bar{\boldsymbol{\tau}} = [\tau_i \quad \tau_d]^T \in \mathbb{R}^2$  como el vector de torques de entrada al robot, siendo  $\tau_i$  y  $\tau_d$  los torques en las ruedas izquierda y derecha, respectivamente. La linealización se desea en el punto de equilibrio  $\mathbf{z}_e = [\mathbf{0}^T \quad \mathbf{0}^T]^T$ , y como los torques de entrada en el punto de equilibrio deben de ser cero, entonces  $\bar{\boldsymbol{\tau}}_e = [0 \quad 0]^T$ . Por lo tanto, es posible obtener un modelo linealizado de la forma

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\bar{\boldsymbol{\tau}} \quad (13)$$

donde las matrices  $\mathbf{A}$  y  $\mathbf{B}$  son definidas como

$$\begin{aligned}
 \mathbf{A} &= \left. \frac{\partial \mathbf{f}(\mathbf{z}, \bar{\boldsymbol{\tau}})}{\partial \mathbf{z}} \right|_{\mathbf{z}=\mathbf{z}_e, \bar{\boldsymbol{\tau}}=\bar{\boldsymbol{\tau}}_e} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & a_1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & a_2 & 0 & 0 & 0 \end{bmatrix}, \\
 \mathbf{B} &= \left. \frac{\partial \mathbf{f}(\mathbf{z}, \bar{\boldsymbol{\tau}})}{\partial \bar{\boldsymbol{\tau}}} \right|_{\mathbf{z}=\mathbf{z}_e, \bar{\boldsymbol{\tau}}=\bar{\boldsymbol{\tau}}_e} = \begin{bmatrix} 0 & 0 & 0 & b_1 & -b_2 & -b_3 \\ 0 & 0 & 0 & b_1 & b_2 & -b_3 \end{bmatrix}^T
 \end{aligned}$$

con

$$\begin{aligned}
 a_1 &= -\frac{(gm_p^2 l_p^2)}{(m_p l_p^2 + I_{py})h - m_p^2 l_p^2} \\
 a_2 &= \frac{gm_p l_p h}{(m_p l_p^2 + I_{py})h - m_p^2 l_p^2} \\
 b_1 &= \frac{m_p l_p^2 + I_{py}}{r((m_p l_p^2 + I_{py})h - m_p^2 l_p^2)} \\
 b_2 &= \frac{L}{r(I_{bz} + I_{pz} + I_{rz} + I_{lz} + L^2(m_l + m_r) + \frac{L^2(I_{ly} + I_{ry})}{r^2})} \\
 b_3 &= \frac{m_p l_p}{r((m_p l_p^2 + I_{py})h - m_p^2 l_p^2)} \\
 h &= (m_b + m_l + m_p + m_r) + \frac{I_{ly}}{r^2} + \frac{I_{ry}}{r^2}
 \end{aligned}$$

Es bien sabido, por la literatura, que un sistema de la forma (13) puede ser estabilizado por un control de realimentación de estados lineal dado por

$$\bar{\boldsymbol{\tau}} = -\mathbf{K}\mathbf{z}. \quad (14)$$

El llamado controlador LQR emplea un algoritmo de optimización bien definido (ver, e.g. [10]) para seleccionar la matriz  $\mathbf{K}$ .

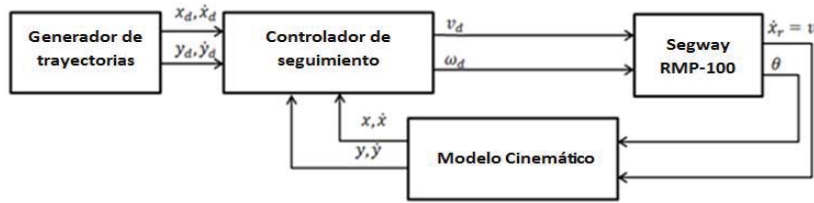


Figura 3: Esquema del control de seguimiento.

Como la salida del LQR son las fuerzas generalizadas de cada rueda (denotadas aquí como  $\tau_i$  y  $\tau_d$ ), y las entradas del robot son  $v_d$  y  $\omega_d$ , se requiere aplicar una transformación. Suponiendo que las masas y momentos de inercia de las ruedas se desprecian es posible escribir

$$F_{x_r} = [m_b + m_p] \ddot{x}_r = \frac{1}{r} [\tau_i + \tau_d]$$

$$\tau_{\theta_r} = I_{bz} + I_{pz} \ddot{\theta} = \frac{L}{r} [\tau_d - \tau_i].$$

de manera que

$$v_d = \dot{x}_{rd} = \frac{1}{r[m_b + m_p]} \int [\tau_d + \tau_i] dt$$

$$\omega_d = \dot{\theta}_d = \frac{L}{r[I_{bz} + I_{pz}]} \int [\tau_d - \tau_i] dt.$$

## V. RESULTADOS EXPERIMENTALES

La Tabla I muestra los parámetros requeridos en el modelo dinámico de la plataforma. Estos parámetros fueron estimados utilizando procedimientos de identificación convencionales para el prototipo.

Tabla I: Parámetros del modelo dinámico del robot

Clave	Descripción	Valor	Unidades
$r$	radio de las ruedas	0.2	m
$L$	distancia media entre ruedas	0.234	m
$l_p$	distancia al c.d.m del péndulo	0.1513	m
$m_b$	masa de la base	0.0	kg
$m_p$	masa del péndulo	53.6825	kg
$m_r$	masa de la rueda derecha	2.35	kg
$m_l$	masa de la rueda izquierda	2.35	kg
$I_{bz}$	m.d.i de la base en el eje Z	0.0	kgm <sup>2</sup>
$I_{py}$	m.d.i del péndulo en el eje Y	3.592	kgm <sup>2</sup>
$I_{pz}$	m.d.i del péndulo en el eje Z	1.8394	kgm <sup>2</sup>
$I_{ly}$	m.d.i de la rueda izquierda en Y	0.0550	kgm <sup>2</sup>
$I_{lz}$	m.d.i de la rueda izquierda en Z	0.0282	kgm <sup>2</sup>
$I_{ry}$	m.d.i de la rueda derecha en Y	0.0550	kgm <sup>2</sup>
$I_{rz}$	m.d.i de la rueda derecha en Z	0.0282	kgm <sup>2</sup>

m.d.i = momento de inercia

### V-A. Controlador de seguimiento

Para este experimento se configuró el robot en modo de autobalanceo, lo que activa el controlador de balanceo interno para el péndulo. La Fig. 3 muestra el diagrama para el control

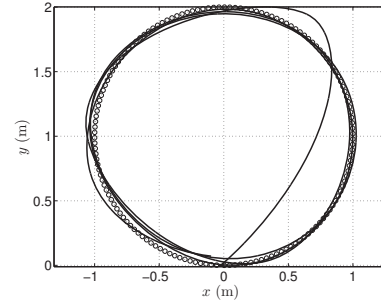


Figura 4: Trayectorias: deseada y actual en el plano X-Y.

de seguimiento. La trayectoria deseada para el control de seguimiento está dada por

$$x_d = c \sin(\omega_0 t) \quad (15)$$

$$y_d = c[1 - \cos(\omega_0 t)] \quad (16)$$

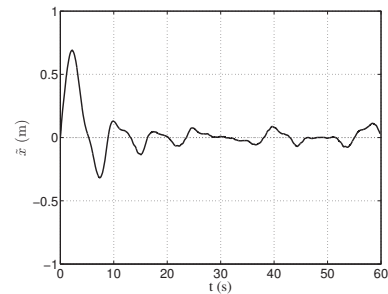
con

$$c = 1.0 \text{ m} \quad \omega_0 = \frac{2\pi \text{ rad}}{15 \text{ s}}.$$

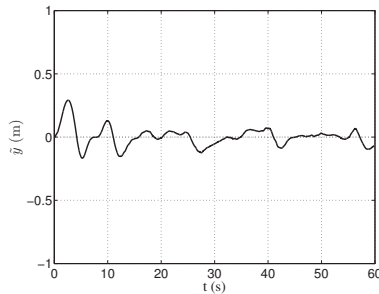
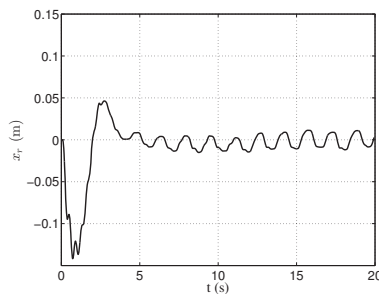
La Fig. 4 muestra tanto la trayectoria deseada (en línea punteada) como la trayectoria real seguida por el robot. Las Figs. 5 y 6 muestran la evolución temporal de  $\tilde{x}$  y  $\tilde{y}$ .

### V-B. Control de balanceo

En este caso el robot se configura en modo tractor (sin el controlador de autobalanceo). La tarea consiste en balancear el péndulo manteniendo la postura inicial del robot (i.e.  $x_r = 0$  y  $\theta_r = 0$ ).


 Figura 5: Evolución temporal de  $\tilde{x}$ .




 Figura 6: Evolución temporal de  $\tilde{y}$ .

 Figura 7: Evolución temporal de  $x_r$ .

Para los experimentos se considera que el robot tiene una inclinación inicial de  $-0.16$  [rad]. Después de una etapa de sintonización se decidió utilizar la siguiente matriz de ganancias

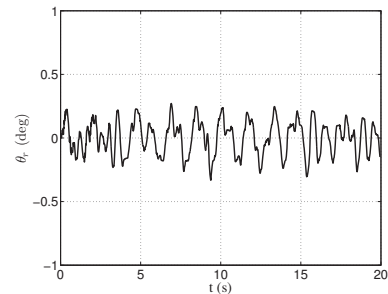
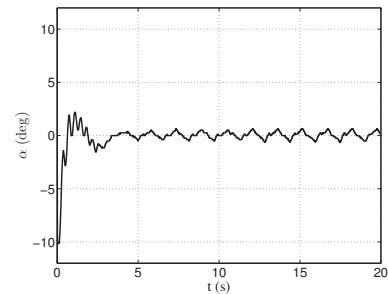
$$K = \begin{bmatrix} -17.77 & -5.1 & -126.37 & -15.73 & -5.2 & -49.2 \\ -17.77 & 5.1 & -126.37 & -15.73 & -5.2 & -49.2 \end{bmatrix}.$$

que aunque no cumple con las condiciones óptimas del LQR sí asegura que los polos del sistema lineal en lazo cerrado estén en el semiplano izquierdo. Cabe mencionar también que, adicionalmente a la ley de control (14) se le tuvo que agregar una acción integral sólo en la parte correspondiente al estado  $x_r$ , para reducir los errores en estado estacionario en la dirección de esta variable.

Las figuras 7, 8 y 9 muestran la evolución temporal de  $x_r$ ,  $\theta_r$  y  $\alpha$ , respectivamente, durante 20 segundos. Nótese que de acuerdo a la Fig. 9, después del transitorio el robot es efectivamente balanceado con inclinaciones alrededor de  $\pm 2$  grados. La orientación se mantiene entre  $\pm 0.6$  grados (Fig. 8), y el desplazamiento lineal  $x_r$  (Fig. 7) se mantiene oscilando entre  $\pm 0.1$  metros gracias a la acción integral.

## VI. CONCLUSIÓN

La plataforma robótica Segway RMP-100 ha sido estudiada en este trabajo. Primero se utilizó la formulación lagrangiana para calcular el modelo dinámico de la plataforma. También se mostró cómo implementar controladores en tiempo real en este robot. Se probaron dos controladores por separado: un controlador lineal LQR para balancear el péndulo y un controlador no lineal para seguimiento de una trayectoria


 Figura 8: Evolución temporal de  $\theta_r$ .

 Figura 9: Evolución temporal de  $\alpha$ .

deseada. Se incluyeron resultados experimentales para ambos controladores.

## AGRADECIMIENTOS

Este trabajo fue apoyado parcialmente por el TecNM y PRODEP.

## REFERENCIAS

- [1] S. W. Nawawi, M.Ñ. Ahmad, and J. H. Osman, "Real-time control of a two-wheeled inverted pendulum mobile robot," in *Proc. of WorldAcademy of Science, Engineering and Technology*, 2008.
- [2] Y.-S. Ha and S. Yuta, "Trajectory tracking control for navigation of the inverse pendulum type self-contained mobile robot," pp. 65–80, 1996.
- [3] F. Grasser, A. D'Arrigo, S. Colombi, and A. C. Rufer, "JOE: A mobile, inverted pendulum," *IEEE Transactions on Industrial Electronics*, vol. 49, no. 1, pp. 107–114, 2002.
- [4] A. Salerno and J. Angeles, "On the nonlinear controllability of a quasiholonomic mobile robot," *2003 IEEE International Conference on Robotics and Automation*, vol. 3, 2003.
- [5] S. W. Nawawi, M.Ñ. Ahmad, and J. H. Osman, "Control of two-wheeled inverted pendulum mobile robot using full order sliding mode control," in *Proc. of International Conference on Man-Machine Systems*, Langkawi, Malaysia, 2006.
- [6] K. Pathak, J. Franch, and S. K. Agrawal, "Velocity and position control of a wheeled inverted pendulum by partial feedback linearization," *IEEE Transactions on Robotics*, vol. 21, no. 3, pp. 505–513, 2005.
- [7] K. D. Do and G. Seet, "Motion control of a two-wheeled mobile vehicle with an inverted pendulum," *J. Intell. Robot. Syst.*, vol. 60, pp. 577–605, 2010.
- [8] L. Sciacivco and B. Siciliano, *Modelling and Control of Robot Manipulators*. Springer Verlag, 2000.
- [9] G. Oriolo, A. D. Luca, and M. Vendittelli, "WMR control via dynamic feedback linearization: design, implementation, and experimental validation," *IEEE Transactions on Control Systems Technology*, vol. 10, no. 6, 2002.
- [10] J. P. Hesphana, *Linear Systems Theory*. Princeton University Press, 2009.

# Control en orientación de un Micro vehículo submarino mediante un sistema embebido

Jesus Monroy

Departamento de Control Automático  
UMI-LAFMIA, CINVESTAV-CNRS  
Cinvestav  
México, D.F., México  
Email: jmonroy@ctrl.cinvestav.mx

Eduardo Campos

UMI-LAFMIA, CINVESTAV-CNRS  
México, D.F., México  
Universidad Politécnica de Pachuca  
Hidalgo, México  
Email: ecampos@ctrl.cinvestav.mx

Jorge Torres

Departamento de Control Automático  
UMI-LAFMIA, CINVESTAV-CNRS  
Cinvestav  
México, D.F., México  
Email: jtorres@ctrl.cinvestav.mx

**Abstract**—En este artículo se presenta el diseño, modelado y control de un micro submarino (Cookie ROV) para realizar tareas de inspección en espacios cerrados, se utiliza un sistema embebido (gumstix) para efectuar la autonomía, además del uso de dispositivos y materiales de bajo costo. El principal objetivo es la implementación y comparación de los controladores PD, PID y PD saturado para la orientación del ( $\mu$ AUV) mediante una IMU. Se describe el modelo dinámico del prototipo mediante las ecuaciones de Euler-Lagrange. Finalmente, se presenta la validación del desempeño del vehículo mediante simulación y resultados experimentales en tiempo real.

**Index Terms**—Micro Vehículo Submarino Autónomo ( $\mu$ AUV), prototipo, controladores PD, PID y PD saturado, control, experimentos en tiempo real.

## I. INTRODUCCIÓN

Actualmente, hay un creciente interés en el desarrollo de vehículos submarinos autónomos (AUV), dado la búsqueda de nuevas fuentes de energía, así como la investigación en desastres ambientales. Hoy en día, se están tratando muchas tareas exclusivas con el uso de robots submarinos en la exploración de aguas profundas. Existen diferentes tipos de vehículos submarinos no tripulados, entre ellos se encuentran los vehículos sumergibles operados de forma remota "vehículos teledirigidos (ROV)" y los "vehículos submarinos autónomos (AUV)", los cuales son utilizados en función de la profundidad del mar y de la misión a realizar.

Hay una gran variedad de aplicaciones en el uso de vehículos no tripulados para evitar tareas que ponen en riesgo al ser humano. La correcta operación de estos vehículos requiere de la convergencia de diferentes disciplinas, entre ellas; el control automático juega un rol importante cuando se trata con dinámicas complejas no lineales, como es el caso de los AUVs. Los AUVs han sido utilizados tradicionalmente para investigación oceanográfica en ambientes que pueden ser considerados como infinitamente grandes. De manera que estos vehículos han sido aplicados para tareas de mapeo y monitoreo en áreas localizadas del océano. Algunas misiones como: la inspección en oleoductos, construcción de infraestructura submarina, estudio de arrecifes y monitoreo de ecosistemas en el mundo marino, son aplicaciones que se encuentran en espacios no reducidos.

Existen otras aplicaciones que pueden beneficiarse del uso de AUVs, tal como el monitoreo de estanques de almacenamiento nuclear, instalaciones de tratamiento de aguas residuales y estudio de ecosistemas en cavernas acuáticas. Estos ambientes difieren del océano en relación a sus dimensiones (estos pueden ser considerados como espacios cerrados) y posibles áreas irregulares. Una clave necesaria para monitoreo de procesos en espacios cerrados, es la habilidad de tomar mediciones en lugares específicos.

Los AUVs tradicionales tienden a ser de gran escala (dimensión en metros) y alto costo, haciendo estos inadecuados para ambientes de pequeña escala. Un interesante tema, en el punto de vista de control automático; es que muy comúnmente los movimientos en el plano horizontal y vertical se encuentran acoplados, es decir, para el impulso de avance se requiere mover el vehículo en el plano vertical. Esto significa que los diseños de AUVs tradicionales son inapropiados para ser usados como plataformas sensoras de pequeña escala ([1]). Las principales características de los micro vehículos, con respecto a vehículos tradicionales, son su tamaño y su maniobrabilidad (ver [2] y [3]).

El prototipo  $\mu$ AUV (Cookie ROV), se presenta en la figura 1, su diseño tiene la forma de un cilindro con un diámetro de 15 cm y un espacio vertical de 12 cm. Los diferentes grados de libertad son actuados por cuatro propulsores de propela. Cabe mencionar que la forma geométrica del prototipo permite que los movimientos verticales y horizontales sean considerados desacoplados. Además de que este prototipo tiene la particular característica de tener estabilidad en lazo abierto en alabeo y cabeceo (ángulos de roll y pitch), debido a la calibración del peso y fuerza de flotabilidad de forma colineal, la cual facilita su geometría.

Este trabajo presenta detalles del desarrollo de la plataforma y la comparación e implementación de estrategias de control PD, PID y PD saturado para la dinámica en guiñada (ángulo yaw). De manera que el estudio se lleva a cabo tanto en simulación como experimentos en tiempo real. La sección II presenta más detalles del diseño y ventajas del  $\mu$ AUV. En la sección III se describe brevemente la dinámica del modelo del prototipo  $\mu$ AUV mediante las ecuaciones de Euler-Lagrange, mientras que en la sección IV se presenta

la implementación de estrategias de control para el ángulo yaw mediante un controladores PD, PID y PD saturado (ver [4]). En la sección V se detalla la simulación de las leyes de control implementadas. Finalmente, se concluye con las observaciones y el trabajo futuro propuesto en la sección VI.



Figure 1: Prototipo experimental Cookie ROV.

## II. DESCRIPCIÓN DEL PROTOTIPO

### A. Sistema Embebido

La arquitectura de hardware del prototipo consiste de un sistema embebido que incluye un gumstix overo fire, el cual está basado en una arquitectura ARM Cortex-A8 de 720 MHz, 512 MB de memoria RAM, WiFi, Bluetooth, expansión a puertos USB, I2C, GPIO, PWM, UART y SPI. Tiene un diseño compacto (20 x 65 mm de dimensiones) con un peso de 5.6 g. El sistema operativo Angstrom está basado en plataforma Linux para ser programado en lenguajes de alto nivel. Este sistema embebido contiene una central inercial (CMPS10) que se comunica con el microprocesador a través del puerto I2C y proporciona los ángulos de euler (roll, pitch, yaw), además de las velocidades angulares. También contiene un sensor de presión (BMP085) para obtener la altura e hidrófonos para sensar señales acústicas. El gumstix procesa la información de los sensores, efectúa la ley de control implementada y envía las señales de control a los actuadores (motores trifásicos sin escobillas) controlados por modulación de ancho de pulso (PWM) a través de los controladores de velocidad electrónicos (ESC de 30A). El acceso remoto al sistema se realiza mediante la comunicación WiFi y cuando el prototipo se introduce al agua, la comunicación se lleva a cabo a través de la interface ethernet. En la figura 2 se muestra un esquema resumido de la arquitectura, los componentes del hardware del vehículo y su interacción.

### B. Diseño y Descripción del movimiento

En la figura 3 se presenta el diseño CAD del micro vehículo con la ayuda del programa SolidsWorks. Su diseño

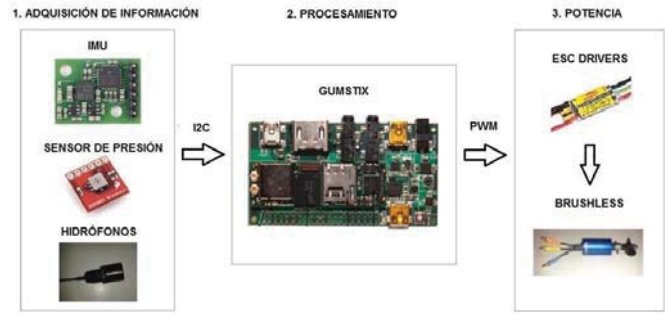


Figure 2: Sistema Embebido del  $\mu$ AUV.

tiene localizados dos motores para el movimiento en el plano horizontal y dos motores para el movimiento en el plano vertical. El referencial  $(O_b, x_b, y_b, z_b)$  corresponde al marco de referencia fijo al cuerpo. El centro  $(O_b)$  de este marco corresponde a el centro de gravedad del vehículo, y sus ejes están alineados con los ejes principales de simetría. El movimiento en el plano horizontal corresponde a la dinámica en *surge* (a lo largo del eje  $x_b$ ) y *sway* (a lo largo del eje  $y_b$ ), mientras que el *heave* representa el movimiento vertical (a lo largo del eje  $z_b$ ). Los ángulos de Euler, alabeo, cabeceo y guiñada denotados por  $(\phi, \theta, \psi)$  describen la orientación del marco de referencia fijo al vehículo con respecto al marco de referencia fijo a la tierra  $(O_I, x_I, y_I, z_I)$ , mientras que  $(x, y, z)$  denotan las coordenadas del centro del marco de referencia fijo al vehículo en el marco referencial de la tierra. El sistema de propulsión consiste de cuatro propulsores, como se muestra en la figura 4, de manera que se genera el movimiento rotacional y traslacional. Con respecto al movimiento de rotación, el movimiento en cabeceo (pitch) es efectuado a través de la diferencia en el control de velocidad de los propulsores 1 y 2. El movimiento en guiñada (yaw) se obtiene de manera similar pero utilizando los propulsores 3 y 4, finalmente el movimiento en alabeo (roll) no está actuado, lo que indica una dinámica desacoplada para este caso, ya que con sólo rotar el vehículo 90 grados en ángulo yaw y activar los motores verticales, se logra obtener el movimiento deseado en el alabeo. Por otro lado, el movimiento traslacional del eje  $z$  es regulado por el incremento y decremento de la combinación en velocidad de los propulsores 1 y 2. Similarmente, los movimientos de traslación a lo largo de los ejes  $x_b$  y  $y_b$  se obtienen mediante el uso de los propulsores 3, 4 y el control del ángulo yaw.

## III. MODELO DINÁMICO

La dinámica del vehículo que se expresa en el marco de referencia fijo al cuerpo, puede escribirse en forma vectorial de acuerdo a [5]:

$$M\dot{\nu} + C(\nu)\nu + D(\nu)\nu + g(\eta) = \tau + w_e \quad (1)$$

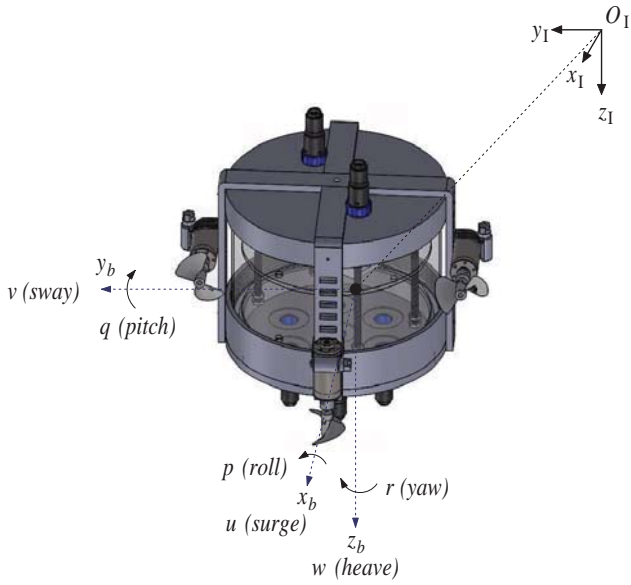


Figure 3: El prototipo  $\mu$ AUV Cookie ROV, con su marco de referencia fijo al cuerpo ( $O_b, x_b, y_b, z_b$ ), y el marco de referencia fijo a la tierra ( $O_I, x_I, y_I, z_I$ ).

$$\dot{\eta} = J(\eta)\nu \quad (2)$$

donde  $M \in \mathbb{R}^{6 \times 6}$  es la matriz de inercia,  $C(\nu) \in \mathbb{R}^{6 \times 6}$  representa la matriz de coriolis de fuerzas centrípetas,  $D(\nu) \in \mathbb{R}^{6 \times 6}$  corresponde a la matriz de amortiguamiento hidrodinámico,  $g(\eta) \in \mathbb{R}^{6 \times 1}$  describe el vector de fuerzas y momentos gravitacional y flotabilidad,  $\tau = (\tau_1, \tau_2)^T = ((\tau_X, \tau_Y, \tau_Z), (\tau_K, \tau_M, \tau_N))^T \in \mathbb{R}^{6 \times 1}$  define el vector de entradas de control;  $w_e \in \mathbb{R}^{6 \times 1}$  representa el vector de perturbaciones;  $\nu = (\nu_1, \nu_2)^T = ((u, v, w), (p, q, r))^T \in \mathbb{R}^{6 \times 1}$  denota el vector de velocidad lineal y angular en el marco de referencia fijo al cuerpo;  $\eta = (\eta_1, \eta_2)^T = ((x, y, z), (\phi, \theta, \psi))^T \in \mathbb{R}^{6 \times 1}$  es el vector de posición y orientación representado en el marco de referencia fijo a la tierra, y  $J(\eta) \in \mathbb{R}^{6 \times 6}$  es la matriz de transformación entre el marco de referencia fijo al cuerpo y el marco de referencia fijo a la tierra (ver más detalles, [6],[7]).

La orientación de vehículos submarinos, relativa al sistema de coordenadas fijo a la tierra, está dado por una transformación de velocidad lineal

$$\dot{\eta}_1 = J_1(\phi, \theta, \psi)\nu_1 \quad (3)$$

donde la secuencia de rotación se escribe como:

$$J_1(\phi, \theta, \psi) = R_\psi R_\theta R_\phi \quad (4)$$

Por otra parte, el vector de velocidad angular fijado al cuerpo  $\nu_2$  y el vector velocidad de Euler  $\eta_2$  están relacionadas a través de una matriz de transformación  $J_2$  como sigue (ver [6])

$$\eta_2 = J_2(\phi, \theta, \psi)\nu_2 \quad (5)$$

las ecuaciones anteriores (4) y (5) pueden escribirse con relación a la notación descrita en [8], como sigue:

$$J(\eta) = \begin{bmatrix} J_1(\eta_2) & 0 \\ 0 & J_2(\eta_2) \end{bmatrix} \quad (6)$$

como consecuencia, esto resulta:

$$\ddot{\eta} = J(\eta)\dot{\nu} + \dot{J}(\eta)\nu \implies \dot{\nu} = J^{-1}(\eta)[\ddot{\eta} - \dot{J}(\eta)J^{-1}(\eta)\dot{\eta}]$$

$$M_\eta(\eta) = J^{-T}(\eta)MJ^{-1}(\eta)$$

$$C_\eta(\nu, \eta) = J^{-T}(\eta)[C(\nu) - MJ^{-1}(\eta)J(\dot{\eta})]J^{-1}(\eta)$$

$$D_\eta(\nu, \eta) = J^{-T}(\eta)D(\nu)J^{-1}(\eta)$$

$$g_\eta(\eta) = J^{-T}(\eta)g(\eta)$$

$$\tau_\eta(\eta) = J^{-T}(\eta)\tau$$

finalmente, la ecuación (1) puede escribirse como

$$M_\eta(\eta)\ddot{\eta} + C_\eta(\nu, \eta)\dot{\eta} + D_\eta(\nu, \eta)\dot{\eta} + g_\eta(\eta) = \tau_\eta(\eta) \quad (7)$$

#### A. Fuerzas y momentos de Gravedad/Flotabilidad

De acuerdo al principio de Arquímedes, la fuerza de flotabilidad  $f_B$  se efectúa en el centro de flotabilidad y actúa en dirección opuesta al peso del vehículo  $f_W$ . Esto conduce a:

$$f_B = - \begin{bmatrix} 0 \\ 0 \\ \rho g \nabla \end{bmatrix} \quad f_W = \begin{bmatrix} 0 \\ 0 \\ mg \end{bmatrix} \quad (8)$$

donde  $\rho$  representa la densidad del fluido,  $g$  la aceleración gravitacional,  $\nabla$  el volumen del fluido desplazado y  $m$  corresponde a la masa del vehículo. Ahora, considerando que  $W = mg$  y  $B = \rho g \nabla$  usando la convención  $zyx$  para la navegación y el control (ver [5]), entonces la matriz de transformación  $J_1(\eta_2) = R_{z,\psi}R_{y,\theta}R_{x,\phi}$  se efectúa para obtener las fuerzas del peso y la flotabilidad respecto al sistema de coordenadas fijo al cuerpo, esto es:

$$F_B = J_1(\eta_2)^{-1}f_B \quad F_W = J_1(\eta_2)^{-1}f_W \quad (9)$$

Por lo tanto:

$$F_B = \begin{bmatrix} B \sin(\theta) \\ -B \cos(\theta) \sin(\phi) \\ -B \cos(\theta) \cos(\phi) \end{bmatrix} \quad F_W = \begin{bmatrix} -W \sin(\theta) \\ W \cos(\theta) \sin(\phi) \\ W \cos(\theta) \cos(\phi) \end{bmatrix} \quad (10)$$

De forma que, las fuerzas de restitución que actúan sobre el vehículo son  $f_g = F_B + F_W$ , esto es:

$$f_g = \begin{bmatrix} (B - W) \sin(\theta) \\ (W - B) \cos(\theta) \sin(\phi) \\ (W - B) \cos(\theta) \cos(\phi) \end{bmatrix} \quad (11)$$

Por otro lado, los momentos de restauración se describen mediante la siguiente ecuación:

$$m_g = r_w \times F_W + r_b \times F_B \quad (12)$$

donde  $\mathbf{r}_w = [x_w, y_w, z_w]^T$  y  $\mathbf{r}_b = [x_b, y_b, z_b]^T$  representan las posiciones del centro de gravedad (CG) y el centro de flotabilidad (CB) respectivamente. Basado en el diseño del vehículo y para reducir aún más el análisis, el origen del marco de referencia fijo al cuerpo se selecciona en el centro de gravedad, esto implica que  $\mathbf{r}_w = [0, 0, 0]^T$ ; mientras que el centro de flotabilidad es  $\mathbf{r}_b = [0, 0, -z_b]^T$ . Para propósitos prácticos, la fuerza de flotabilidad debe ser mayor que el peso, es decir,  $W - B = -f_b$ . Observe que  $f_b$  debe ser menor que la fuerza producida por los propulsores. Entonces de las ecuaciones (11) y (12), se tiene:

$$\mathbf{g}(\boldsymbol{\eta}) = \begin{bmatrix} \mathbf{f}_g \\ \mathbf{m}_g \end{bmatrix} = \begin{bmatrix} f_b \sin(\theta) \\ -f_b \cos(\theta) \sin(\phi) \\ -f_b \cos(\theta) \cos(\phi) \\ -z_b B \cos(\theta) \sin(\phi) \\ -z_b B \sin(\theta) \\ 0 \end{bmatrix} \quad (13)$$

### B. Fuerzas y momentos generados por los propulsores

La figura 4 muestra las fuerzas generadas por los propulsores actuando sobre el micro submarino, éstas son descritas relativamente en el sistema coordenado fijo al cuerpo, como:

$$\hat{\mathbf{f}}_1 = \begin{bmatrix} 0 \\ 0 \\ f_1 \end{bmatrix}; \hat{\mathbf{f}}_2 = \begin{bmatrix} 0 \\ 0 \\ f_2 \end{bmatrix}; \hat{\mathbf{f}}_3 = \begin{bmatrix} f_3 \\ 0 \\ 0 \end{bmatrix}; \hat{\mathbf{f}}_4 = \begin{bmatrix} f_4 \\ 0 \\ 0 \end{bmatrix}$$

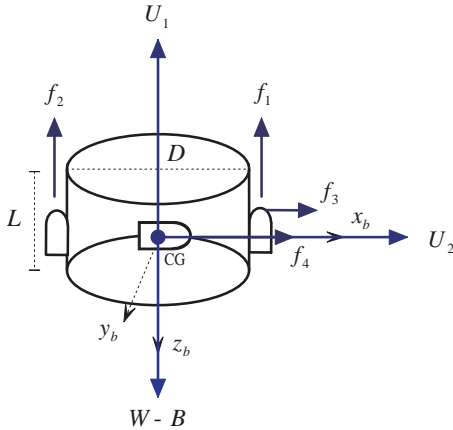


Figure 4: Fuerzas generadas por los cuatro propulsores del vehículo  $f_i \forall i = [1...4]$ . CG describe la posición del centro de gravedad.

Las fuerzas generadas por los propulsores que actúan en el micro submarino  $\boldsymbol{\tau}_{u_1}$  (ver Figura 4) son descritas como sigue:

$$\boldsymbol{\tau}_1 = \begin{bmatrix} U_2 \\ 0 \\ U_1 \end{bmatrix} = \begin{bmatrix} \sum_{i=3}^4 \hat{f}_i \\ 0 \\ \sum_{i=1}^2 \hat{f}_i \end{bmatrix} \quad (14)$$

Al resumir y utilizar la notación de [8], se tiene que:

$$\boldsymbol{\tau}_1 = \begin{bmatrix} \tau_X \\ \tau_Y \\ \tau_Z \end{bmatrix} = \begin{bmatrix} f_3 + f_4 \\ 0 \\ f_1 + f_2 \end{bmatrix} \quad (15)$$

y los momentos fijos al cuerpo generados por las fuerzas arriba presentadas, se definen como:

$$\boldsymbol{\tau}_2 = \sum_{i=1}^4 \mathbf{l}_i \times \hat{\mathbf{f}}_i \quad (16)$$

donde  $\mathbf{l}_i = (l_{ix}, l_{iy}, l_{iz})$  es el vector de posición de la fuerza  $\hat{\mathbf{f}}_i \forall i = 1, \dots, 4$ , con respecto al marco de referencia fijo al cuerpo. Entonces los momentos generados por los propulsores se describen como:

$$\boldsymbol{\tau}_2 = \begin{bmatrix} \tau_K \\ \tau_M \\ \tau_N \end{bmatrix} = \begin{bmatrix} 0 \\ l_{1x}f_1 + l_{2x}f_2 \\ l_{3y}f_3 + l_{4y}f_4 \end{bmatrix} \quad (17)$$

por lo tanto:

$$\boldsymbol{\tau} = \begin{bmatrix} f_3 + f_4 \\ 0 \\ f_1 + f_2 \\ 0 \\ l_{1x}f_1 + l_{2x}f_2 \\ l_{3y}f_3 + l_{4y}f_4 \end{bmatrix} \quad (18)$$

## IV. ESTRATEGIA DE CONTROL

Con el fin de estabilizar el vehículo frente a posibles perturbaciones, se propone una ley de control basado en el uso de funciones de saturación. Las funciones de saturación limitan las señales de control agresivas enviadas a los actuadores. Se omite el análisis de estabilidad, el lector interesado puede ver más detalles en [9]. Debido a la geometría del vehículo y la baja latencia de operación, se supone que las matrices constantes  $M, C, D, g$ , son conocidas. Entonces, aplicando la siguiente entrada de control,

$$\boldsymbol{\tau}_\eta(\boldsymbol{\eta}) = C_\eta(\boldsymbol{\nu}, \boldsymbol{\eta})\dot{\boldsymbol{\eta}} + D_\eta(\boldsymbol{\nu}, \boldsymbol{\eta})\boldsymbol{\eta} + g_\eta(\boldsymbol{\eta}) + M_\eta(\boldsymbol{\eta})\boldsymbol{\tau} \quad (19)$$

al sistema (7), y cancelando los factores en común, se tiene:

$$M_\eta(\boldsymbol{\eta})\ddot{\boldsymbol{\eta}} = M_\eta(\boldsymbol{\eta})\boldsymbol{\tau} \quad (20)$$

donde:

$$\boldsymbol{\tau} = J^{-T}(\boldsymbol{\eta}) \begin{bmatrix} \boldsymbol{\tau}_1 \\ \boldsymbol{\tau}_2 \end{bmatrix} \quad (21)$$

en consecuencia, se produce:

$$\begin{aligned} \ddot{x} &= u_2 c_\theta c_\psi + u_1 [c_\phi c_\psi s_\theta + s_\phi s_\psi] \\ \ddot{y} &= u_2 c_\theta s_\psi + u_1 [c_\phi s_\theta s_\psi - c_\psi s_\phi] \\ \ddot{z} &= -u_2 s_\theta + u_1 c_\theta c_\phi - w_p \\ \ddot{\phi} &= \tau_\phi \\ \ddot{\theta} &= \tau_\theta c_\phi - \tau_\psi s_\phi \\ \ddot{\psi} &= -\tau_\phi s_\theta + \tau_\theta c_\theta s_\phi + \tau_\psi c_\phi c_\theta \end{aligned} \quad (22)$$

### A. Control PD saturado

Para el diseño del controlador, es común asumir que los parámetros hidrodinámicos involucrados en el modelo dinámico del vehículo submarino son desconocidos. En efecto, estos dependen de las condiciones y propiedades difíciles de modelar o estimar, como la masa agregada al vehículo, la fricción de superficie, generación de vortex, características del

fuido, etc. Sin embargo, se propone utilizar un controlador PD saturado.

Sea  $u(t)$  un controlador PD el cual se describe a continuación:

$$u(t) = k_1 e(t) + k_2 \frac{de(t)}{dt} \quad (23)$$

Donde  $e(t) = r(t) - y(t)$  es el error,  $r(t)$  representa la referencia,  $y(t)$  es la salida medida, y  $(k_1, k_2)$  son las ganancias proporcional y derivativa. En la ecuación 23 es notable que si  $e(t) \rightarrow \infty$ , entonces  $u(t) \rightarrow \infty$ ; esto podría generar oscilaciones en el sistema o de otro modo saturar los actuadores. Con el objetivo de prevenir daño en los actuadores, se propone utilizar una función de saturación en cada término de la ecuación 23.

$$\begin{aligned} \tau_\phi &= 0 \\ \tau_\theta &= f(\theta, \dot{\theta}) = -\sigma_{b\theta_2}(k_{\theta_2}\dot{\theta}) - \sigma_{b\theta_1}(k_{\theta_1}\theta) \\ \tau_\psi &= f(\psi, \dot{\psi}) = -\sigma_{b\psi_2}(k_{\psi_2}\dot{\psi}) - \sigma_{b\psi_1}(k_{\psi_1}\psi) \end{aligned} \quad (24)$$

esto implica que existe un tiempo tal que  $\phi \rightarrow 0$ ,  $\theta \rightarrow 0$  y  $\psi \rightarrow 0$  con lo cual se tiene:

$$\begin{aligned} \ddot{x} &= U_2 & \ddot{\phi} &= 0 \\ \ddot{y} &= 0 & \ddot{\theta} &= f(\theta, \dot{\theta}) \\ \ddot{z} &= U_1 C_\phi - W_p & \ddot{\psi} &= f(\psi, \dot{\psi}) \end{aligned} \quad (25)$$

Se puede notar que el movimiento de traslación sobre el eje  $x$  ahora depende únicamente de la entrada de control  $U_2$  y el movimiento sobre el eje  $z$  depende de  $U_1$ , por lo cual se proponen las siguientes leyes de control:

$$U_1 = \frac{f(z, \dot{z}) + W_p}{C_\phi} \quad U_2 = f(x, \dot{x}) \quad (26)$$

donde

$$\begin{aligned} f(z, \dot{z}) &= -\sigma_{bz_2}(k_{z_2}\dot{z}) - \sigma_{bz_1}(k_{z_1}(z - z_d)) \\ f(x, \dot{x}) &= -\sigma_{bx_2}(k_{x_2}\dot{x}) - \sigma_{bx_1}(k_{x_1}x) \end{aligned} \quad (27)$$

donde las  $k_{ij}$  son constantes positivas tal que los polinomios  $s^2 + k_{ij}s + k_{ij-1}$  sean estables (donde  $i = \theta, \psi, x, z, j = 2$  y  $\sigma_b$  es una función saturación de la forma:

$$\sigma_b(s) = \begin{cases} b & \text{si } s > b \\ s & \text{si } -b \leq s \leq b \\ -b & \text{si } s < -b \end{cases} \quad (28)$$

Sustituyendo las ecuaciones(26) en el sistema(25) se asume que existe un tiempo tal que:

$$\begin{aligned} \ddot{x} &= f(x, \dot{x}) & \ddot{\phi} &= 0 \\ \ddot{y} &= 0 & \ddot{\theta} &= f(\theta, \dot{\theta}) \\ \ddot{z} &= f(z, \dot{z}) & \ddot{\psi} &= f(\psi, \dot{\psi}) \end{aligned} \quad (29)$$

La estrategia de control estabiliza la orientación y posición. Por lo tanto, la ley de control que estabiliza la orientación se propone como:

$$\begin{aligned} \tau_\theta &= \frac{-\sigma_{b\theta_2}(k_{\theta_2}\dot{\theta}) - \sigma_{b\theta_1}(k_{\theta_1}\theta) + \tau_\psi s_\phi}{c_\phi} \\ \tau_\psi &= \frac{-\sigma_{b\psi_2}(k_{\psi_2}\dot{\psi}) - \sigma_{b\psi_1}(k_{\psi_1}\psi) - \tau_\theta s_\phi}{c_\phi} \end{aligned} \quad (30)$$

## B. Control PID

Para un controlador PID, la ley de control se expresa de la siguiente manera:

$$\tau_{PID} = \mathbf{J}(\nu)^T \left[ \mathbf{K}_p \tilde{\eta} + \mathbf{K}_d \dot{\tilde{\eta}} + \mathbf{K}_i \int \tilde{\eta}(\sigma) d\sigma \right] \quad (31)$$

donde  $\tilde{\eta} = \eta_d - \eta$  es el error entre la posición real y la deseada,  $\mathbf{K}_p$  es la matriz de ganancias proporcionales,  $\mathbf{K}_d$  corresponde a la matriz de ganancias derivativas, y  $\mathbf{K}_i$  es la matriz de ganancias integrales.

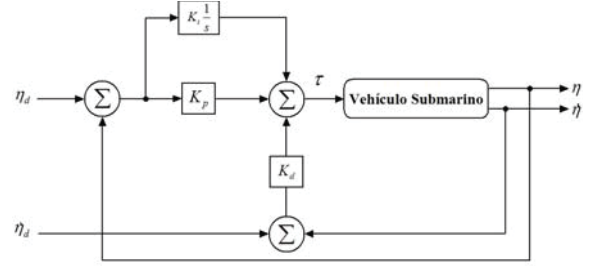


Figure 5: Controlador PID.

Al aplicar un controlador PID clásico para el control en yaw. La entrada de control, se representa como:

$$\tau_\psi = -K_p(\psi - \psi_d) - K_i \int_0^t (\psi - \psi_d) dt - K_d(\dot{\psi} - \dot{\psi}_d) \quad (32)$$

donde  $\tau_\psi$  es la entrada en Newtons para aplicarse alrededor de la rotación del eje  $z$ ,  $K_p$ ,  $K_d$  y  $K_i$  son las ganancias constantes positivas, correspondientes a la ganancia proporcional, derivativa e integral.  $\psi$  es el ángulo yaw medido, mientras que  $\psi_d$  es el ángulo yaw deseado,  $\dot{\psi}$  y  $\dot{\psi}_d$  son sus respectivas derivadas.

La entrada de control de la ecuación 32 se basa en la descripción matemática del controlador PID, dado por:

$$u(t) = K_p \left[ e(t) + \frac{1}{T_i} \int_0^t e(t) dt + T_d \frac{de(t)}{dt} \right] \quad (33)$$

donde  $e(t)$  es la señal de error,  $K_p$  la ganancia proporcional,  $T_i$  el tiempo integral y  $T_d$  el tiempo derivativo.

## V. SIMULACIÓN

Se utiliza Simulink de MatLab para probar las leyes de control seleccionadas en la estabilidad de la posición angular yaw sobre las ecuaciones del modelo dinámico.

Se presentan los resultados de la simulación con el objetivo de verificar el desempeño de la dinámica en yaw, utilizando las leyes de control PD, PID y PD saturado para realizar una comparación y seleccionar la que mejor desempeño efectúe en el sistema. En la figura 11 se muestra la respuesta del sistema al implementar un controlador PD y someterlo a perturbaciones del tipo escalón (ver figura 6). Las ganancias correspondientes equivalen a;  $K_p=0.8$  y  $K_d=0.5$ , mientras que en el gráfico de la figura 7 se presenta la respuesta

del sistema mediante un controlador PID, con la entrada de control  $\tau_{\psi}$  (ver figura 8), cuyas ganancias equivalen a:  $K_p=1.8$ ,  $K_d=1.5$  y  $K_i=1.2$ , es notable que se puede mejorar la estabilidad a través de la sintonización. Para el caso del controlador PD saturado, éste fue sintonizado de igual manera para observar el comportamiento frente a perturbaciones, la evolución del tiempo en la posición angular (yaw) se muestra en la figura (9) y la entrada de control  $\tau_{\psi}$  correspondiente (ver figura 10), las ganancias equivalen a;  $K_p=1.5$  y  $K_d=0.8$ . Observe que para mejorar la estabilidad, se puede sintonizar con más precisión, considerando la característica que tiene este controlador, de permitir saturar la entrada de control y proteger a los actuadores. Los resultados experimentales se presentan en las imágenes (12, 13 y 14), donde es notable que el controlador PD saturado brinda una mejor estabilidad. Cabe mencionar que éste fue sintonizado para conseguir el mejor desempeño.

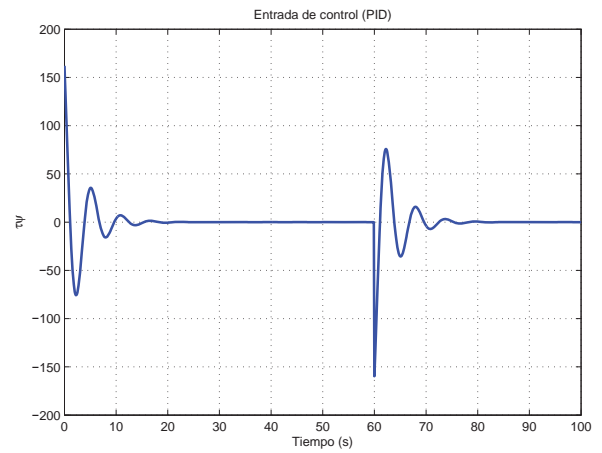


Figure 8: Entrada de control  $\tau_{\psi}$  para el control PID.

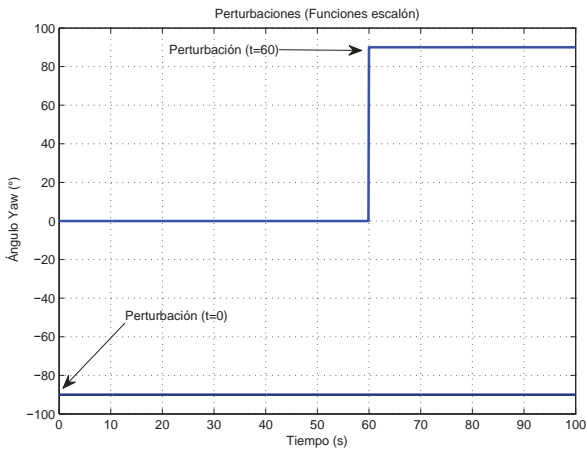


Figure 6: Perturbaciones aplicadas al sistema.

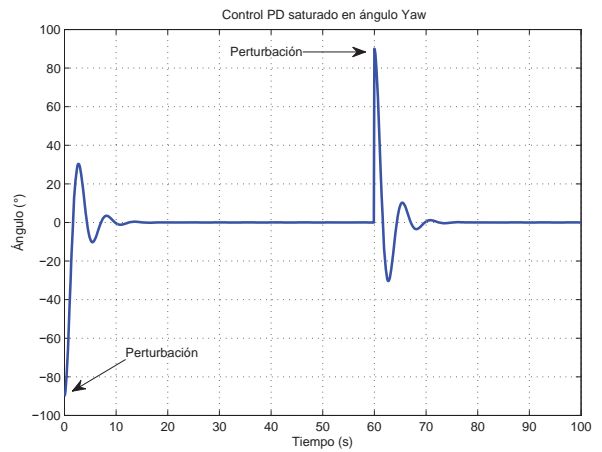


Figure 9: Control PD saturado en yaw (simulación).

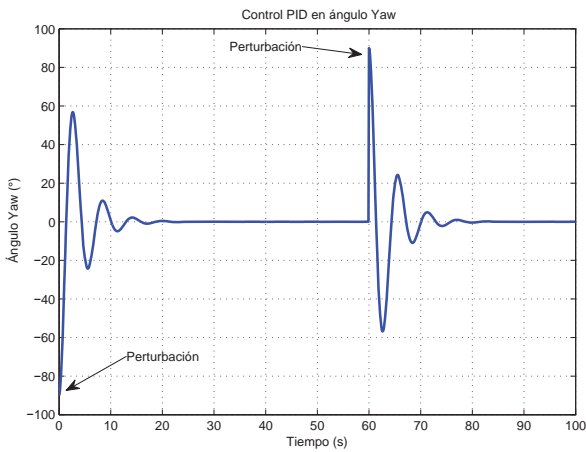


Figure 7: Control PID en yaw (simulación).

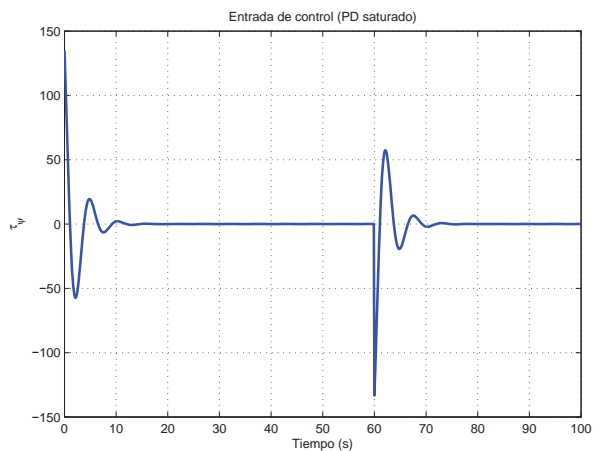


Figure 10: Entrada de control  $\tau_{\psi}$  para el control PD saturado.

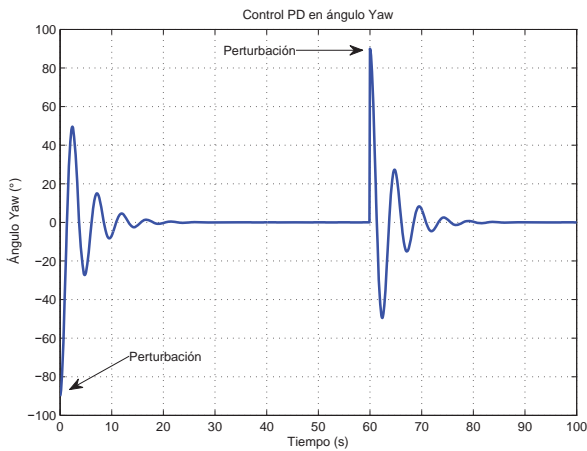


Figure 11: Control PD en yaw (simulación).

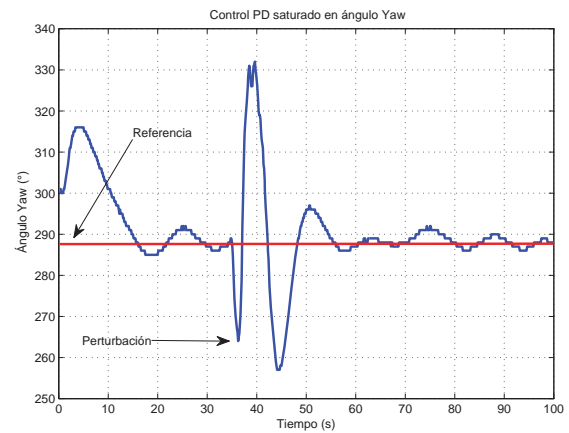


Figure 14: Sintonización experimental del Control PD saturado (yaw).

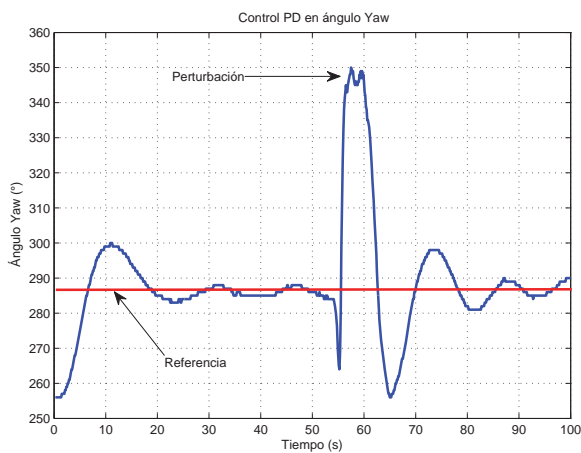


Figure 12: Sintonización experimental del Control PD (yaw).

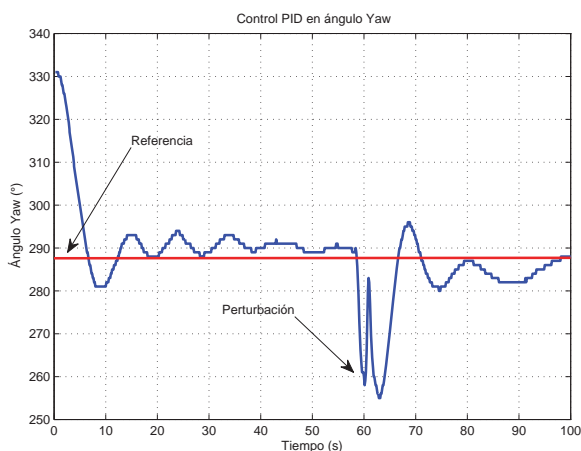


Figure 13: Sintonización experimental del Control PID (yaw).

## VI. CONCLUSIONES Y TRABAJO FUTURO

El desarrollo de este prototipo surgió por la necesidad de tener un vehículo autónomo de pequeña escala para operar en ambientes cerrados y realizar maniobras complejas, el sistema de navegación embebida se redujo tanto como fue posible. La dinámica en los movimientos en el plano horizontal y vertical se consideran desacoplados como clave del problema. Por un lado, en este trabajo inicial se considera el problema de regulación para la orientación en yaw, y por otro lado se esta buscando realizar tareas de cooperación utilizando la visión artificial como trabajo futuro (ver [10]). En este escrito se utilizaron controladores típicos, donde el controlador PD saturado presentó una mejor respuesta. El comportamiento deseado fue validado a través de la simulación en MatLab.

## REFERENCES

- [1] P. Rodríguez, J. Piera *Mini AUV, a platform for future use on marine research for the Spanish Research Council?* Instrumentation ViewPoint 2005.
- [2] Watson Simon A. and Green Peter N. (2014) *Depth Control for Micro-autonomous Underwater Vehicles ( $\mu$ AUVs): Simulation and Experimentation.* International Journal of Advanced Robotic Systems, DOI: 10.5772/57334.
- [3] Sascha Fechner, Jochen Kerdels, Jan Albiez, and Frank (2007) *Design of a  $\mu$ AUV.* Conference: Proceedings of the 4th International AMiRE Symposium (AMiRE-2007).
- [4] Antonelli, G., Fossen, T. I., and Yoerger, D. R. (2008). *Underwater Robotics.* In Siciliano, B. and Khatib, O., editors, Springer Handbook of Robotics, chapter 44. Spring Berlin Heidelberg, Berlin, Heidelberg.
- [5] Thor I. Fossen. *Marine control systems guidance, navigation, and control of shipd, rigs and underwater vehicles.* Marine Cybernetics, 2002.
- [6] Thor H. Goldstein, C.P. Poole and J.L. Safko. *Classical Mechanics.* Addison Wesley Series in Physics, Addison-Wesley, USA, second edition, 1983.
- [7] Marsden, J.E. *Elementary Classical Analysis.* W.H. Freeman and Company, San Francisco, 1974.
- [8] The Society of Naval Architects and Marine Engineers. Nomenclature for Treating the Motion of a Submerged Body Through a Fluid. IN: *Technical and Research Bulletin* No. 1-5.
- [9] A. R. Teel. *Global stabilization and restricted tracking for multiple integrators with bounded controls.* Systems and Control Letters, 1992.
- [10] U.R. Zimmer (2006) *The Australian National University - Information Engineering - Serafina Website.* The Australian National University. [Online]. Available: <http://serafina.anu.edu.au/>



# Diseño de control para evasión de obstáculos en un robot móvil omnidireccional utilizando Programación Genética

Roberto C. Pérez, Eddie Clemente\*, Marlen Meza-Sánchez, E. Bugarin y A. Y. Aguilar-Bustos  
 Tecnológico Nacional de México. Instituto Tecnológico de Ensenada  
 Blvd. Tecnológico 150, Ex-Ejido Chapultepec, C.P. 22750, Ensenada, B. C., México  
 \*eclemente@ite.edu.mx

**Resumen**—En este trabajo se aborda el problema de navegación autónoma de un robot móvil omnidireccional con evasión de obstáculos en un ambiente estático. La propuesta presenta el diseño de un controlador que combina un enfoque de solución clásico con una acción para la evasión del obstáculo obtenida mediante Programación Genética. Se presentan resultados de simulación para varios escenarios en donde se obtienen velocidades acotadas, lo cual garantiza su implementación en un ambiente real.

## I. INTRODUCCIÓN

La navegación autónoma de robots corresponde a un problema importante en el área de la Robótica Móvil. Entre los tipos de robots existentes, los robots móviles omnidireccionales ofrecen la capacidad de realizar cualquier movimiento instantáneo en cualquier dirección sin importar su orientación. Esta característica permite el desarrollo de tareas en donde es necesaria una gran maniobrabilidad como es el caso de la navegación autónoma con evasión de obstáculos.

Un enfoque clásico de solución al problema de navegación autónoma se fundamenta en un proceso iterativo de varias etapas. La primera etapa, consiste en la generación de modelos matemáticos para representar la dinámica de los sistemas robóticos utilizando las leyes de la física. Una vez que la etapa de modelado es concluida, se realiza un análisis de las características del sistema para desarrollar propuestas de algoritmos de control. Estas propuestas se fundamentan en la aplicación de herramientas matemáticas para garantizar propiedades como estabilidad bajo la ley de control y el cumplimiento del objetivo establecido. Una vez terminado el análisis y la comprobación teórica, el controlador deberá ser sintonizado para su implementación y validación experimental.

Ahora bien, la Robótica Evolutiva es un área de investigación cuyo objetivo es la aplicación de técnicas de computación evolutiva (por ejemplo, Algoritmos Genéticos, Programación Genética, Co-evolución, Estrategias Evolutivas, entre otros) para el diseño e implementación de aplicaciones sobre robots, que se inspiran en el principio darwiniano de la reproducción selectiva del más apto [1]. Los algoritmos desarrollados utilizando esta vertiente, simulan el proceso natural de evolución en el cual los miembros de una población cambian sus estructuras para adaptarse al ambiente con la finalidad de sobrevivir. En particular, la Programación Genética se ha utilizado para generar programas en lenguajes de alto nivel que permiten

la navegación de un robot en ambientes restringidos [2]. Como también su implementación en lenguaje ensamblador en pequeños robots, logrando un comportamiento reactivo en los ambiente en que interactúan [3], [4]. De la misma manera se han hecho desarrollos para adaptar la morfología de un robot para tareas específicas, [5], [6]. Otro punto de vista de la robótica evolutiva, combina los algoritmos evolutivos para desarrollar controladores difusos [7], [8]; neurocontroladores [9], [10]; o una combinación de estos [11].

Respecto al uso del enfoque clásico de solución en la Robótica Evolutiva, en [12], se presenta una propuesta para planeación de trayectorias con evasión de obstáculos utilizando lógica difusa para diseñar campos potenciales en ambientes estáticos y con obstáculos en movimiento. El diseño de controladores difusos donde su estructura esta basada en la teoría de Lyapunov y PSO se detalla en [13].

De esta manera, en el presente trabajo se expone una propuesta para resolver el problema de control de movimiento para un robot móvil omnidireccional, que incluye una optimización aplicando Programación Genética, para evasión de obstáculos. Además, los resultados obtenidos muestran que se generan trayectorias con velocidades acotadas y alcanzables para un robot móvil, lo cual garantiza su implementación en un escenario real de experimentación.

El presente documento está organizado como se describe a continuación. En la Sección II se plantea el problema de control para navegación autónoma con evasión de obstáculos para un robot móvil omnidireccional. El diseño de control propuesto utilizando la teoría de control y Programación Genética se presenta en la Sección III. Resultados de simulación del esquema de control propuesto, donde se compara en desempeño con un control propuesto por [14], se muestran en la Sección IV. Finalmente, las conclusiones y trabajo futuro de esta propuesta se presentan en la Sección V.

## II. DESCRIPCIÓN DEL PROBLEMA

Considérese el problema de control para navegación autónoma de un robot móvil omnidireccional de acuerdo al escenario presentado en la Figura 1. El movimiento del robot con respecto a un marco fijo XY se describe por un vector de posiciones  $q = [x \ y]^T$ .

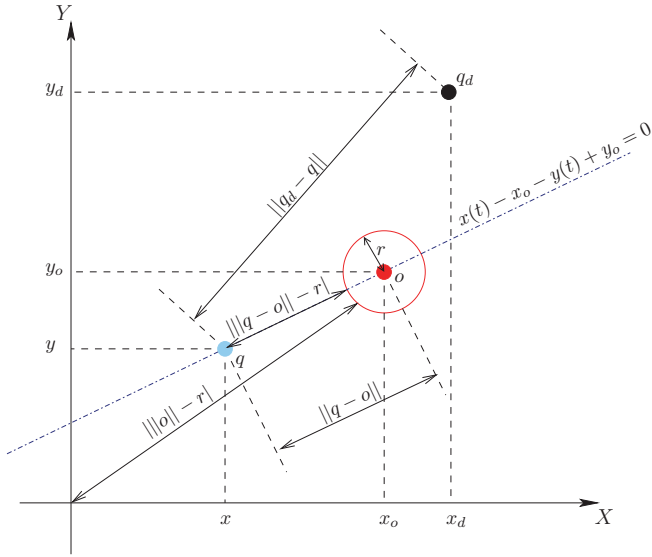


Figura 1. Esquema de configuración del problema de control.

El objetivo de control se define entonces como el diseño de un controlador que lleve asintóticamente al robot de la posición  $q = [x \ y]^T$  a la posición deseada y constante  $q_d = [x_d \ y_d]^T$ ; esto es,

$$\lim_{t \rightarrow \infty} q(t) = q_d, \quad (1)$$

que a su vez, tenga un procedimiento para la evasión de un obstáculo localizado en  $o = [x_o \ y_o]^T$  de radio  $r$ .

El modelo cinemático del robot móvil omnidireccional se describe por

$$\dot{q} = u, \quad (2)$$

donde  $\dot{q} = [\dot{x} \ \dot{y}]^T$ , y  $u = [u_x \ u_y]^T$  describe al vector de velocidades lineales de entradas de control para el robot con respecto a los ejes  $X$  y  $Y$ , respectivamente.

### III. DISEÑO DEL CONTROLADOR

Considere el modelo cinemático descrito en (2) para representar el movimiento del robot móvil omnidireccional. Se propone una ley de control compuesta por

$$u = u_a + u_r, \quad (3)$$

donde  $u_a = K\tilde{q}$  corresponde a una acción de atracción encargada de la convergencia del robot a la posición deseada  $q_d$ , con  $K \in \mathbb{R}^{2 \times 2}$  diagonal definida positiva y  $\tilde{q} = [\tilde{x} \ \tilde{y}]^T = [x_d - x \ y_d - y]^T$ ; mientras que  $u_r = [u_{rx} \ u_{ry}]^T = v[\tilde{y} \ -\tilde{x}]^T$  describe a una acción de repulsión para la evasión del obstáculo localizado en  $o$ , con  $v$  un función escalar por definir y la cual será encontrada mediante un algoritmo evolutivo basado en Programación Genética.

Cabe destacar que este controlador está inspirado en el propuesto por [14], en donde se trabaja con una función  $v$  descrita por (con una ligera variación)

$$v = \frac{\tanh(\lambda[y - L(x)][x_d - x_o])}{\|q - o_q\| + \epsilon}, \quad (4)$$

donde  $\lambda > 0$ ,  $\epsilon > 0$ ,  $L(x)$  es la ecuación de la línea que une el punto de la posición deseada con el centro del obstáculo y  $o_q = o + r \frac{q-o}{\|q-o\|}$ .

#### III-A. Prueba

En esta apartado se realizará la demostración del cumplimiento del objetivo de control (1), en el robot (2) y con la ley de control (3). De manera que la ecuación en lazo cerrado en términos de  $\tilde{q}$  es

$$\dot{\tilde{q}} = -K\tilde{q} - v[\tilde{y} \ -\tilde{x}]^T,$$

en donde se ve claro que el origen es un punto de equilibrio. Ahora considérese la siguiente función candidata de Lyapunov

$$V(\tilde{q}) = \frac{1}{2}\tilde{x}^2 + \frac{1}{2}\tilde{y}^2.$$

La derivada temporal de esta función a lo largo de las trayectorias del sistema en lazo cerrado resulta

$$\dot{V}(\tilde{q}) = -\tilde{x}^2 - \tilde{y}^2 = -\tilde{q}^T \tilde{q} < 0.$$

De esta manera se demuestra que el origen del sistema en lazo cerrado es asintóticamente estable, significando que el objetivo de control (1) es satisfecho (indistintamente de  $v$ ).

#### III-B. Diseño de la función de repulsión utilizando Programación Genética

La programación genética es un algoritmo de búsqueda dirigido, el cual se inspira en la teoría de la evolución natural de Darwin, para descubrir un programa de computadora o función matemática, que sea capaz de resolver una tarea bien definida. Un elemento importante en este proceso es la función objetivo o criterio de búsqueda, el cual indica cuantitativamente el rendimiento de alguna solución propuesta. En este caso, el rendimiento de cada solución debe ponderar que la velocidad alcanzada por el robot sea acotada dentro de un rango bien definido, penalizar las colisiones del robot contra el obstáculo y premiar las soluciones que se aproximen a la distancia más corta entre la posición inicial del robot y la meta. Así la función de rendimiento  $f_i$  se define como sigue:

$$f_i = \frac{1}{2} \cdot (e^{-\frac{1}{2} \cdot (do_i - dr_i)^2}) + \frac{1}{2} \cdot (e^{-\frac{1}{2} \cdot (v_o - v_{max_i})^2}); \quad (5)$$

donde  $do_i$  es la distancia más corta posible que el robot puede recorrer, definida como  $do_i = \|q(0) - q_d\|$ ;  $dr_i$  es la magnitud de la trayectoria recorrida por el robot;  $v_{max}$  es la velocidad máxima alcanzada por el robot y  $v_{maxd}$  es la velocidad máxima deseada. Nótese que  $dr_i$  y  $v_{max}$  quedan acotadas a través de la aplicación de una función Gaussiana centrada en  $do_i$  y  $v_o$ ; donde ambas funciones Gaussianas se ponderan por un factor de  $\frac{1}{2}$ , tal que, el valor máximo de  $f_i$  es la unidad. De esta manera, si se consideran  $n$  escenarios de prueba, entonces la función objetivo,  $F(i)$ , del algoritmo evolutivo se define como:

$$F(i) = \left(\frac{1}{n} \cdot \sum_{i=1}^n f_i\right) \cdot \frac{1}{n_k + 1} \cdot \frac{1}{n_t + 1}; \quad (6)$$

donde,  $n_k$  es el número de colisiones, y  $n_t$  es el número de trayectorias donde no fue alcanzada la meta. Así, la función objetivo es un promedio del rendimiento de la trayectoria que siguió el robot en cada uno de los  $n$  escenarios de prueba. Entonces, el promedio es ponderado por el número de colisiones y la frecuencia en que no se alcanzó la posición deseada. De esta manera, el algoritmo evolutivo selecciona las mejores soluciones a partir de las cuales construye una nueva generación con soluciones diferentes. De forma que, tras varias generaciones se maximiza el valor de la función  $F(i)$ .

Otra parte clave en la implementación del algoritmo evolutivo es la definición de las funciones y terminales de los árboles sintácticos que representan a cada solución. Éstas se han elegido debido a su uso frecuente en el diseño de controladores típicos. El cuadro I, muestra las expresiones propuestas.

Cuadro I. FUNCIONES Y TERMINALES.

	Expresión	Descripción	
Funciones	$+, -, \cdot, /$	Funciones aritméticas	
	$\cos, \sen, \tan, \csc, \sec, \cot$	Funciones trigonométricas	
	$\cosh, \sinh, \tanh, \operatorname{csch}, \operatorname{sech}, \operatorname{coth}$	Funciones hiperbólicas	
	$\arccos, \operatorname{arcsen}$	Funciones trig. inversas	
	$\log, \exp$	Funciones trascendentales	
	$\sqrt{\cdot}, 2^{\cdot}$	Funciones exponenciales	
	$\max(), \min()$	Función máximo y mínimo	
	$\operatorname{norm}(), \operatorname{abs}()$	Norma y valor absoluto	
	Terminales	$x, y$	Posición en $x$ y $y$ del robot
		$x_d, y_d$	Posición deseada en $x$ y $y$
$x_o, y_o, r$		Posición en $x, y$ y radio del obstáculo	
$  q - q_d  ,   q - o  $		Distancia del robot a la meta y al obstáculo	
$   o   - r ,    q - o   - r $		Frontera del obstáculo y distancia del robot a la frontera del obstáculo	
$\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon)$		Commutación de dirección del robot dependiendo de la posición del obstáculo	

IV. RESULTADOS DE SIMULACIÓN

Se propusieron 10 diferentes escenarios con 15 condiciones iniciales distintas. En este caso se consideran como escenarios diferentes cuando la posición del obstáculo varía entre ellos. Así, cada resultado propuesto por la programación genética fue probado en este conjunto de escenarios. Los parámetros empleados durante la evolución son descritos en el cuadro II. Mientras que el parámetro de diseño propuesto para  $v_{maxd} = 2$  y la ganancia  $K$  es igual a la matriz identidad.

Cuadro II. PARÁMETROS UTILIZADOS PARA LA PROGRAMACIÓN GENÉTICA.

Tamaño de la población	100 individuos
Número de generaciones	50 generaciones
Probabilidad de mutación	0.2
Probabilidad de cruzamiento	0.8
Método de inicialización	mitad y mitad
Método de selección	lexicográfico

El comportamiento de desempeño de las soluciones descubiertas a lo largo de las generaciones, se muestra en la figura 2. En este caso, el valor máximo del desempeño es

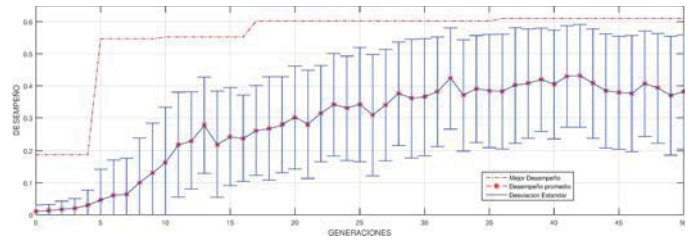


Figura 2. Comportamiento de la evolución a lo largo de 50 generaciones.

Cuadro III. SOLUCIONES DESCUBIERTAS A LO LARGO DE LA EVOLUCIÓN.

Solución	Expresión $v$	Desempeño
1	$\operatorname{asin}(\operatorname{csch}\left(\frac{\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon)}{\operatorname{asin}(\cos(\tanh( q - o  - r))}\right))$	0.6097
2	$\operatorname{csch}\left(\frac{\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon)}{\operatorname{asin}(\cos(\tanh( q - o  - r))}\right)$	0.6021
3	$\operatorname{csch}\left(\frac{\cot(\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon))}{\operatorname{asin}(\cos(\tanh( q - o  - r))}\right)$	0.5520
4	$\operatorname{coth}(\cot\left(\frac{\operatorname{csch}(x_d)}{\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon)}\right))$	0.5465

la unidad. Nótese que las soluciones obtenidas con un valor mayor a 0.5 implican que han alcanzado la meta en un tiempo prácticamente finito y sin colisionar con el obstáculo. De esta forma, se han obtenido las soluciones que se muestran en el cuadro III. Las cuales cumplen con los valores de diseño. Cabe señalar que estas soluciones utilizan la expresión  $\operatorname{sign}(x(t) - x_o - y(t) + y_o + \epsilon)$ , lo que produce un cambio de signo o dirección en la funciones descubiertas. De igual manera, las soluciones con un mayor valor en desempeño utilizan la expresión  $\tanh(|q - o| - r)$ , el cual es un valor ponderado de la distancia del robot al límite del obstáculo. También, es notable el uso de funciones trigonométricas circulares e hiperbólicas.

De las soluciones descritas en el cuadro III, se ha seleccionado la de mejor desempeño para mostrar su comportamiento en los diez escenarios propuestos. La figura 3 muestra las velocidades desarrolladas por el robot en cada una de las trayectorias trazadas en los diferentes escenarios. Cabe señalar que estas velocidades en todos los casos son cercanas al valor deseado de 2. También es notable que los cambios de dirección en velocidad son generalmente suaves, donde solo en los primeros 4 escenarios muestran una pequeña serie de conmutaciones en dirección mientras el robot se aproxima a la meta. La figura 4 expone las rutas seguidas por el robot. Nótese que en ningún caso existe colisión con el obstáculo, donde su tendencia es seguir una trayectoria suave hasta el objetivo.

Comparaciones con lo propuesto por [14], bajo los mismos escenarios, se muestran en las figuras 5 y 6; donde la magnitud de velocidades alcanzadas en los escenarios 5, 6, 8, 9 y 10 se encuentran en el intervalo de las decenas de metros por segundo, ver figura 5. De igual manera, al aproximarse a la meta la velocidad tiende a conmutar de signo produciendo un efecto oscilatorio en la trayectoria, ver figura 6. Con lo cual, aplicando la métrica de desempeño propuesta, esta solución tiene un resultado de 0.1789, aún cuando el robot llegue a la meta y no existan colisiones.

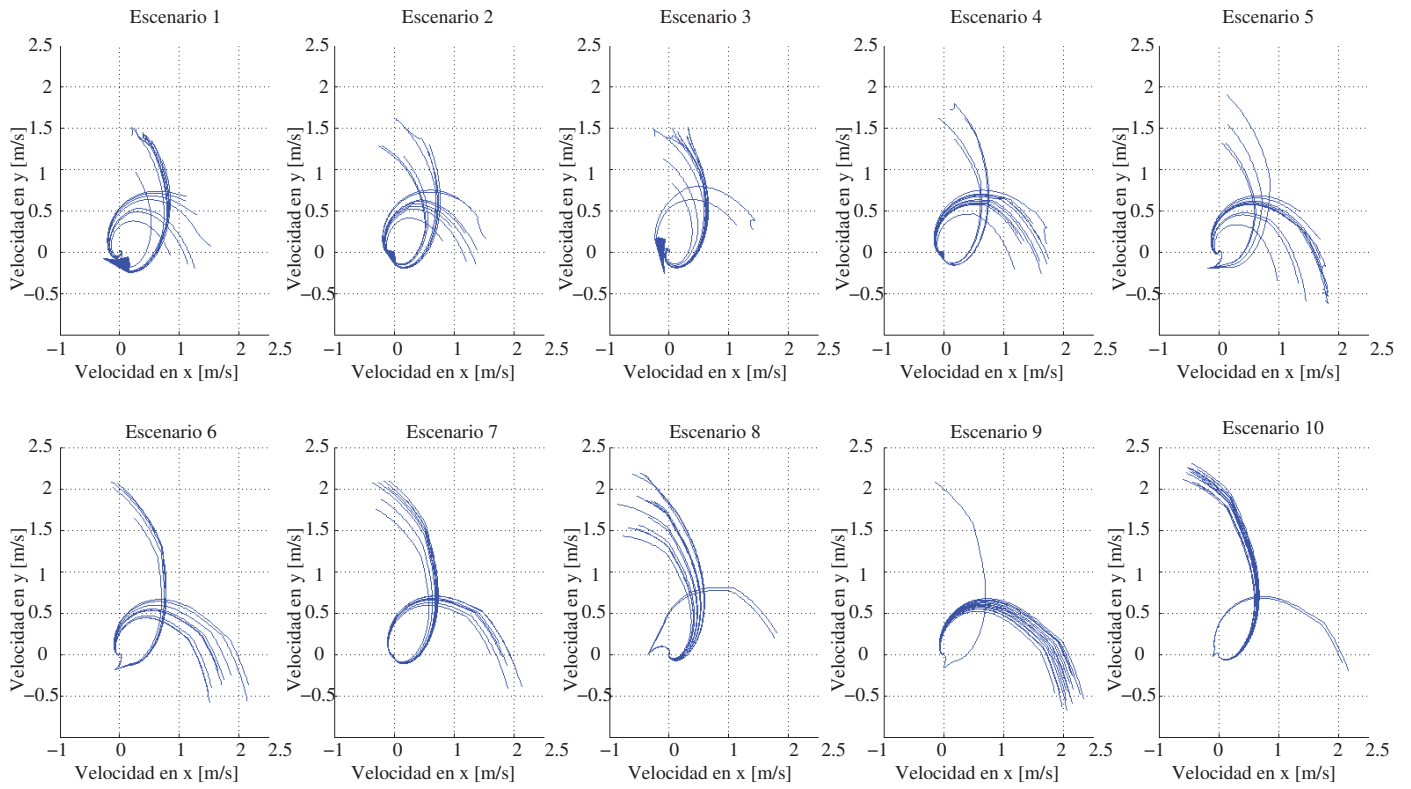


Figura 3. Velocidades resultantes al emplear la mejor solución dada por el cómputo evolutivo, en cada uno de los 10 escenarios propuestos.

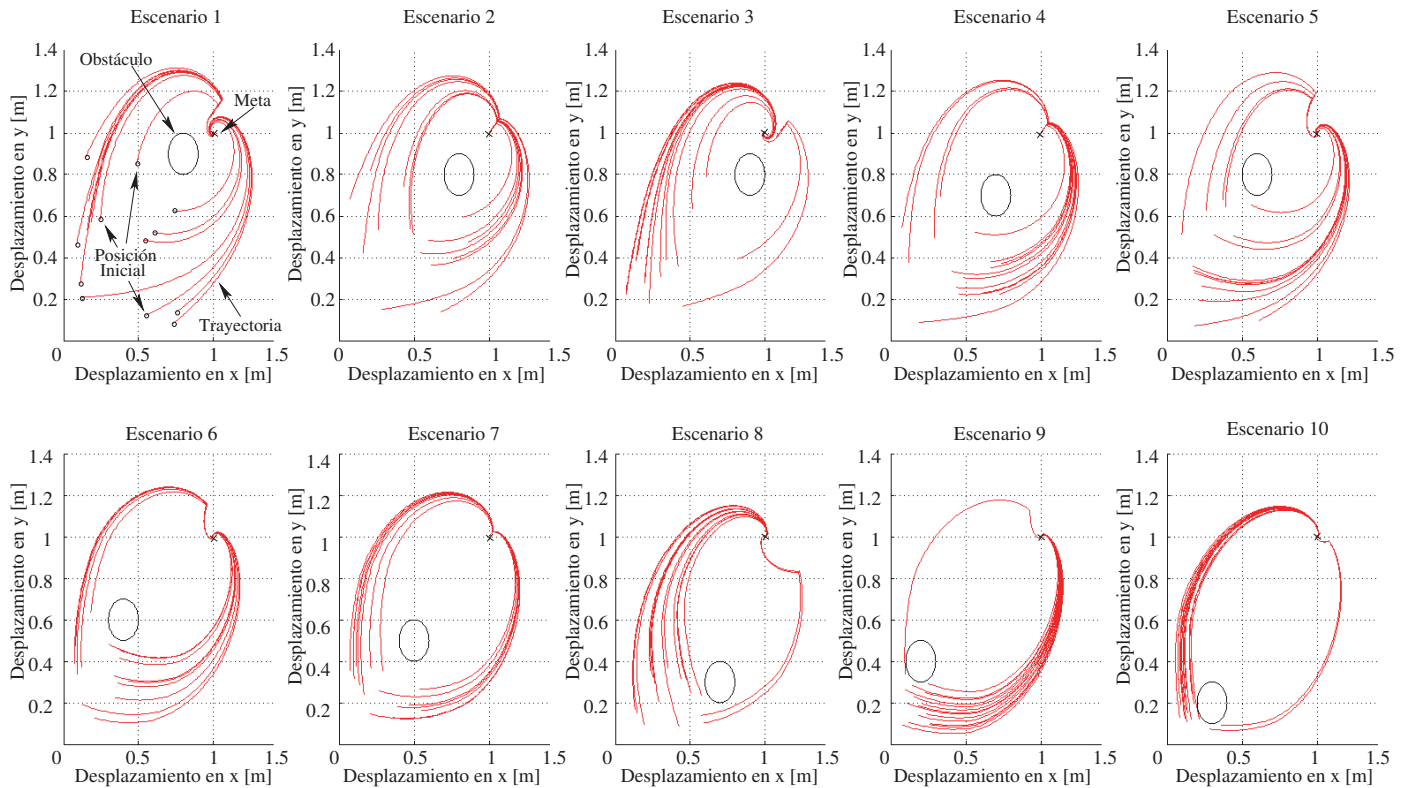


Figura 4. Trayectorias generadas utilizando la mejor solución dada por el cómputo evolutivo, en cada uno de los 10 escenarios propuestos.

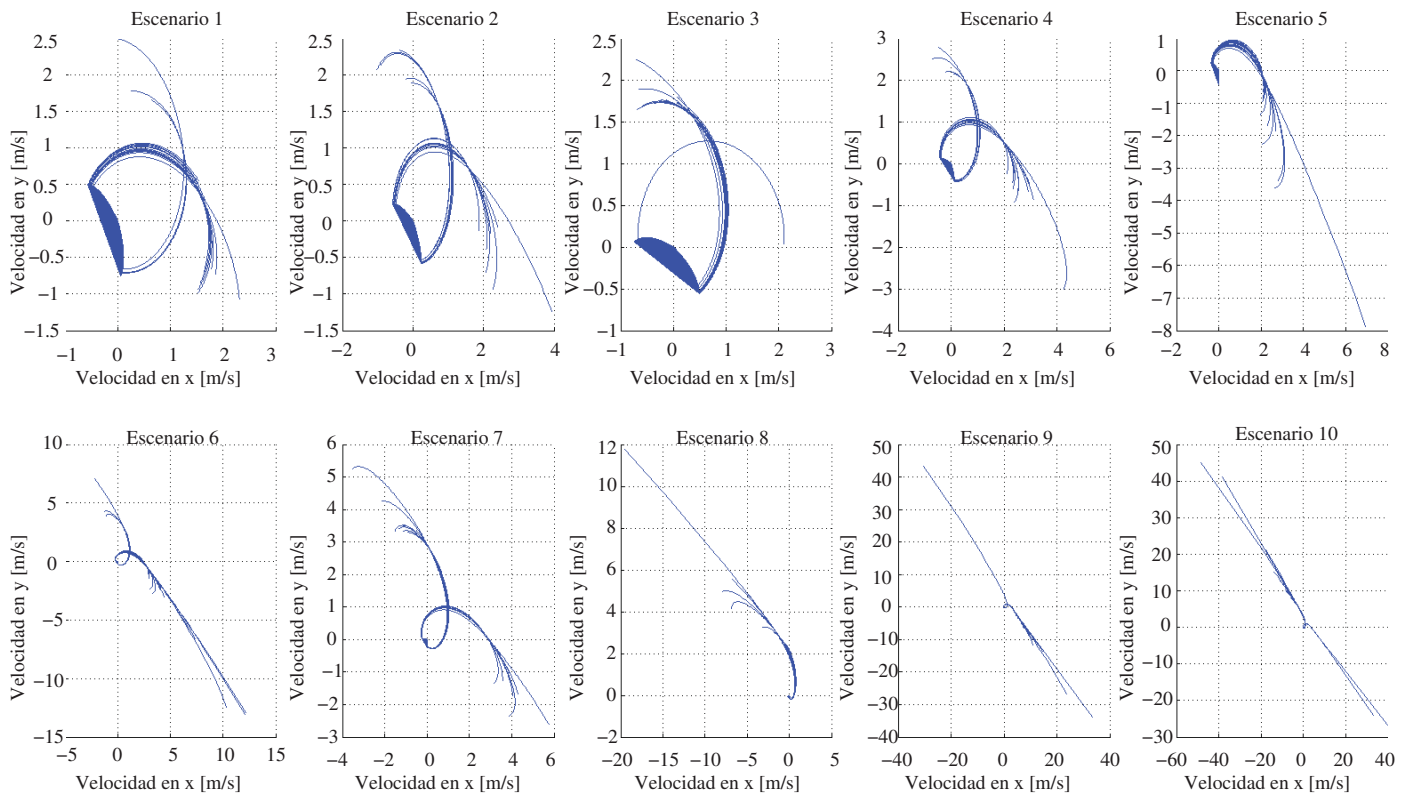


Figura 5. Velocidades generadas al aplicar el modelo propuesto por Kamel, et. al. [14], en los 10 escenarios propuestos

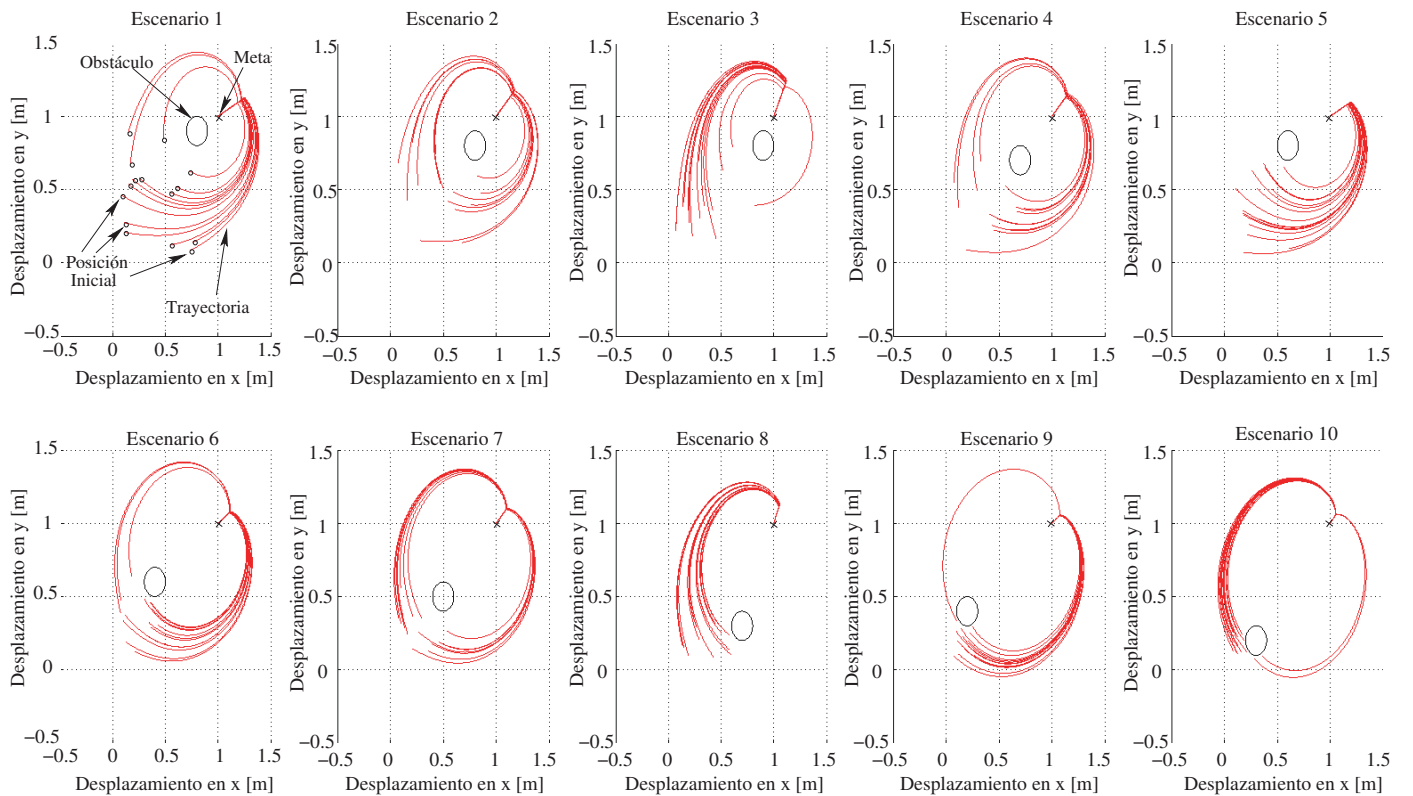


Figura 6. Trayectorias resultantes de la aplicación del modelo propuesto por Kamel, et. al. [14], en los 10 escenarios propuestos.

La comparación entre la solución 1 obtenida y el desarrollo propuesto por [14] se describe en la tabla IV, donde se expone la velocidad máxima alcanzada y la magnitud de la trayectoria mínima, seguida por el robot. Nótese que en todos los casos, la velocidad máxima desarrollada al aplicar la solución propuesta por la programación genética, se encuentra cercana al parámetro de diseño de 2 en velocidad, en contraparte, las magnitudes de las velocidades obtenidas al aplicar el modelo de [14] son mayores. Lo cual representaría un problema de diseño al llevar su implementación en un ambiente real. Más aún, al aplicar la solución encontrada, la distancia recorrida por el robot es menor.

Cuadro IV. COMPARACIÓN ENTRE LA MEJOR SOLUCIÓN Y EL MODELO PROPUESTO POR KAMEL, ET. AL, [14].

Escenario	Velocidad máxima		Trayectoria mínima	
	Kamel et. al, [14].	Solución 1	Kamel et. al, [14].	Solución 1
1	2.54	1.53	1.91	0.89
2	4.13	1.62	1.64	0.82
3	2.35	1.53	1.84	0.86
4	5.20	1.78	1.72	1.10
5	10.49	1.93	1.56	0.94
6	17.79	2.21	1.76	1.25
7	6.33	2.17	1.78	1.28
8	22.76	2.25	1.68	1.26
9	53.00	2.41	1.91	1.43
10	66.00	2.34	2.00	1.52

## V. CONCLUSIÓN Y TRABAJO FUTURO

El presente trabajo muestra una metodología para el diseño de un control para la evasión de obstáculos en un ambiente estático, donde se combina un enfoque de solución clásico y la programación genética. Con lo cual, se obtuvo un diseño que muestra una velocidad acotada y un menor recorrido por el robot. La mejor solución obtenida se compara con la propuesta hecha por Kamel, et. al. [14] notándose su buen desempeño. Un trabajo futuro consiste en realizar una experimentación exhaustiva y llegar a implementaciones con robots reales. Así como desarrollar modelos que impliquen obstáculos o metas en movimiento.

## AGRADECIMIENTOS

Se agradece el apoyo financiero del Tecnológico Nacional de México, del Instituto Tecnológico de Ensenada, del PRO-DEP y del CONACYT (Proyecto 166654 y Proyecto Cátedras 2459).

## REFERENCIAS

- [1] E. Haasdijk, N. Bredeche, S. Nolfi, A. E. Eiben, "Evolutionary Robotics", *Evol. Intel.*, vol. 7, pp. 69 – 70, Aug. 2014.
- [2] J. Koza, *Genetic Programming on the Programming of Computers by Means of Natural Selection*. Cambridge, MA: MIT Press, 1992.
- [3] P. Nordin, W. Banzhaf, "An On-Line Method to Evolve Behavior and to Control a Miniature Robot in Real Time with Genetic Programming". *Adapt. Behav.*, vol. 5, no. 2, pp. 107–140, Jan. 1997.
- [4] M. Ebner, "Evolution of a Control Architecture for a Mobile Robot". *Proceedings of the Second International Conference on Evolvable Systems*, Switzerland, 1998, pp. 303–310.
- [5] H. H. Lund, J. Hallam and Wei-Po Lee, "Evolving robot morphology", *Evolutionary Computation*, 1997., IEEE International Conference on, Indianapolis, IN, 1997, pp. 197–202.

- [6] L. Brodbeck, S. Hauser, F. Iida, "Morphological Evolution of Physical Robots through Model-Free Phenotype Development". *PLoS ONE*, vol. 10, no. 6, June 2015.
- [7] M. Algabri, H. Mathkour, H. Ramdane and M. Alsulaiman, "Comparative study of soft computing techniques for mobile robot navigation in an unknown environment". *Comput. Hum. Behav.*, vol. 50, pp. 42–56, Sept. 2015.
- [8] M. F. Hamza, H. J. Yap, I. A. Choudhury, "Recent advances on the use of meta-heuristic optimization algorithms to optimize the type-2 fuzzy logic systems in intelligent control". *Neural Comput. Appl.*, pp. 1–21, Nov. 2015.
- [9] D. Floreano, J. Urzelai, "Evolutionary robots with online self-organization and behavioral fitness", *Neural Networks*, vol. 13, no. 4–5, pp. 431–443, May-June 2000.
- [10] L. Trujillo, G. Olague, E. Lutton, F.F. De Vega, L. Dozal, E. Clemente, "Speciation in behavioral space for evolutionary robotics". *J. Intell. Robot. Syst.* vol. 64, no. 3-4, pp. 323–351, Jan 2011.
- [11] J. K. Pothal, D. R. Parhi, "Navigation of multiple mobile robots in a highly clutter terrains using adaptive neuro-fuzzy inference system", *Robot. Auton. Syst.*, vol. 72, pp. 48–58, 2015,
- [12] M.A.K. Jaradat, M.H. Garibeh, E.A. Feilat, "Autonomous mobile robot dynamic motion planning using hybrid fuzzy potential field", *Soft Comput.*, vol. 16, no. 1, pp. 153 – 164, June 2012.
- [13] K.D. Sharma, A. Chatterjee, A. Rakshit, "A PSO-Lyapunov Hybrid Stable Adaptive Fuzzy Tracking Control Approach for Vision-Based Robot Navigation", *IEEE Trans. Instrum. Meas.*, vol. 61, no. 7, pp. 1908 – 1914, 2012.
- [14] M. A. El Kamel, L. Beji, A. Abichou and S. Mammari, "A novel obstacle avoidance approach for mobile robot system including target capturing". In *Proc. of the 2nd Mediterranean Conference on Intelligent Systems and Automation (CISA 2009)*, Tunisia, 2009, pp. 249–253.

# Path follower Algorithm for a Nao Humanoid Robot

Martín Montes Rivera, Marving O. Aguilar Justo

Dept. Postgraduate Studies and Research  
Universidad Politécnica de Aguascalientes  
Aguascalientes, Mex.

[martin.montes, marving.omar]@upa.edu.mx

**Abstract**— This paper proposes a path follower algorithm, programmed on a NAO humanoid robot using monocular camera vision. The proposed algorithm requires a Sobel filter, as well as a Canny filter for edge detection, with erode and dilate operations for noise suppress and uses Hough lines transform for detecting the path. The resulting algorithm is programmed using Python 2.7 and OpenCV 2.4.9 library.

**Keywords**— path follower, computer vision, humanoid robot, robotics

## I. INTRODUCTION

The word robot was first time used by Karel Capek in 1921 referring to mechanical beings or humanoid robots, which resemble humans [1, 2].

Movement algorithms of mechanical humanoids like Karel Capek robot, require great computational power for process kinematic and dynamic models, but their movement is not the only requirement [2].

Powerful computers are also needed to use artificial intelligence in humanoid robots for implementing algorithms in mechanical beings to solve different situations, and in the future get emotions, consciousness, coexist with humans, and make them easier their life [1, 2, 3].

Nowadays humanoid robots are a reality although they do not exhibit the behaviour shown in science fiction, but researchers are improving their algorithms due to the increasing of computational power, which have brought new possibilities and robots that are becoming like the Karel Capek concept [4, 5].

Aldebaran is the company that builds Nao (Fig. 1), a humanoid robot with 58 cm height, two cameras, two ultrasonic sonars, two bumpers and an infrared sensor, which help the humanoid robot to make many tasks according to the implemented algorithms [2, 6, 7, 8].



Fig. 1. Nao humanoid robot of Aldebaran [7].

Mediatec company sells Aldebaran products in Mexico and, with the objective of improving their robot distribution and relationship with engineering students, this company have done during the past two years the Nao contest. In this event there are individual, team and exhibition categories; one of the individual events is to run following a path [2, 6, 8].

This paper shows our proposed algorithm to follow a path with the characteristics of the run event in Robotics and Artificial Intelligence Nao 2015 contest, the proposed algorithm is tested using default walk function of Nao robot and the design of the running algorithm will be future work.

## II. BACKGROUNDS

Mobile robots following a desire trajectory, a desired object or a path represented with visible or non-visible lines are common situations in the robotics movement problem [9].

There have been different approaches to solve the problem of a robot following a desired trajectory.

The works in [10] and [11] show how to implement line followers, which are robots that follow a desired path marked with a line using infrared sensors to track it, if it is visible or magnetic sensors if it is not.

Sometimes robots following a specific object is required, the algorithm for a robot leader follower, where a robot follows another robot and the state of leader robot is estimated with a relative equation in discrete domain is proposed in [12]. The proposed algorithm in [13] allows to move several robots following a leader in straight paths.

Another algorithm for track objects is the shown in [14], where a human follower using monocular vision to detect an image in the back of the person followed is implemented in a wheel robot.

When there are obstacles while a robot it is following a path it is required to avoid them and fuzzy logic is a common tool for robot movement and navigation with obstacle avoidance, like the control strategy in [15].

Other works use different sensors to detect and follow a path like in [16] where a human follower is implemented with detection based on the Microsoft Kinect platform.

An algorithm implemented for path tracking in a Nao humanoid robot is shown in [17], where the robot follows the

path according to a sequence of target images. But this imply the previous knowledge of the path.

Our algorithm allows a Nao robot to track a straight path without any previous knowledge like is required in [17]. This work is designed to satisfy the contest specifications in [8] (Fig. 2), using a single camera and image processing. Taking into account to face the problem that the nature of the mechanical behaviour of the NAO humanoid robot exhibits, which is that the orientation of the robot changes continuously and randomly when is walking or running.

### III. CONTEST GENERAL ESPECIFICATIONS

Nao run event imply that the humanoid robot follows a desired path between two lines as fast as it is possible. Departure line is black with 0.18 m length, finish line is red with 0.3 m length, path lines are also black and has 0.025 m width and 0.7 m separated each other [8].

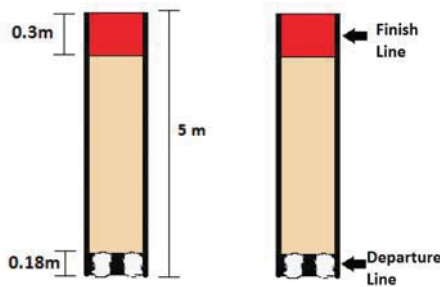


Fig. 2. Path to follow according to contest specs [8].

The robot that finishes first without going out of the path will win the run [8].

### IV. CONCEPTUAL FRAMEWORK

Nao has several sensors but since path is marked with painted lines; it is not possible to sense them with other tool different from the Nao robot cameras.

#### A. Hough line transform

A vision processing tool commonly used for detecting lines, circles and ellipsoids in digital images is the Hough transform introduced by Paul Hough in 1962 [18, 19].

The Hough algorithm obtain the lines in the edges of a picture by getting all the coordinate points in the edge for a Cartesian map  $(x, y)$  like shown in Fig. 3.

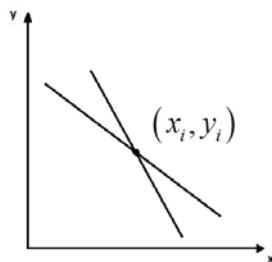


Fig. 3. Lines in a digitalized image that cross each other in  $(x_i, y_i)$ .

Since all lines cross the  $(x_i, y_i)$  point, then all satisfy the equation (1).

$$y_i = mx_i + c \tag{1}$$

First Hough algorithm change equation (1) to equation (2) for calculating the values of  $m$  and  $c$  that matches with the line coordinates in Cartesian map  $(x, y)$ .

$$c = y_i - mx_i \tag{2}$$

Equation (2) generates a line in the map  $(c, m)$  (Fig. 4).

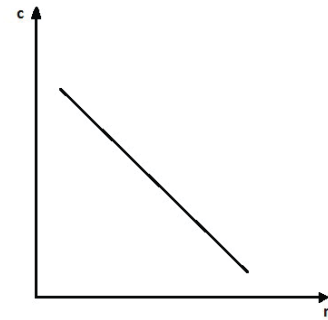


Fig. 4. Points  $(x, y)$  plotted in  $(c, m)$  map.

Hough for line detection initially propose different pairs  $(c, m)$ , and verify the number of points  $(x, y)$  that match with the proposed values, then with a threshold line is confirmed when enough points are detected in the line produced with the proposed  $(c, m)$ .

An improvement of Hough algorithm line is detected with the line polar form like in equation (3) and Fig. 5, which avoid problems when  $m$  is not defined because two points  $a$  and  $b$  of a line are equals or  $x_a - x_b = 0$ .

$$r = x \cos(\theta) + y \sin(\theta) \tag{3}$$

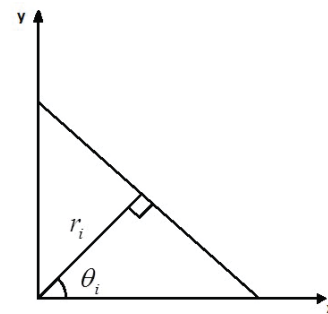


Fig. 5. Polar representation of a line.

With the polar representation several pairs  $(r, \theta)$  are proposed and the number of coincidences of pairs  $(x, y)$ , reflect the lines in the digitalized image.

In polar form the curve for pairs  $(r, \theta)$  is shown in the Fig. 6, where the coincident points are straight lines detected.



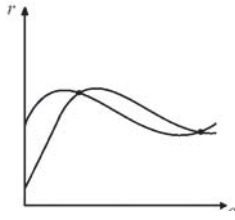


Fig. 6. Map variation of chosen polar values  $(r, \theta)$ .

In general, the Hough transform algorithm proposes  $(r, \theta)$  values while evaluates the number of coincidences for the  $(x, y)$  points in the digitalized image, the most coincidences the great probability to detected line [19].

### B. Sobel Filter

Implementation of Hough transform for detecting lines require a detecting edges filter, with the possibility of avoid horizontal lines like the ones in finish and departure.

Detecting edge filter developed in 1968 by Sobel and Feldman receives the name of Sobel Operator, as many other edge detectors assume that there is a discontinuity in the intensity function when the edge appears [20].

The edge and discontinuity are detected by the gradient calculus obtained by equation (4)

$$G(x, y) = \sqrt{\left[\frac{\partial f}{\partial x}\right]^2 + \left[\frac{\partial f}{\partial y}\right]^2} \approx \left|\frac{\partial f}{\partial x}\right| + \left|\frac{\partial f}{\partial y}\right| \quad (4)$$

A particularity in Sobel Operator is that it is possible to separate horizontal and vertical edges by separating the partial derivatives [20].

Partial derivatives can be calculated by equation (5) and equation (6) for horizontal and vertical separation, respectively.

$$\left|\frac{\partial f}{\partial x}\right| = \frac{1}{k+2} \left[ (w_{1,3} + k \cdot w_{2,3} + w_{3,3}) - (w_{1,1} + k \cdot w_{2,1} + w_{3,1}) \right] \quad (5)$$

$$\left|\frac{\partial f}{\partial y}\right| = \frac{1}{k+2} \left[ (w_{1,1} + k \cdot w_{1,2} + w_{1,3}) - (w_{3,1} + k \cdot w_{3,2} + w_{3,3}) \right] \quad (6)$$

With  $k = 2$  for Sobel Operator and  $w_{i,j}$  are the elements of the convolution matrix used to get the derivatives like is shown in equations (7) and (8).

$$\left|\frac{\partial f}{\partial x}\right| \rightarrow \begin{pmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & 1 \end{pmatrix} \quad (7)$$

$$\left|\frac{\partial f}{\partial y}\right| \rightarrow \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix} \quad (8)$$

### C. Canny Filter

The edge detector proposed by John F. Canny in 1986 is composed filter that has low lever error, minimal response and the good localization or distance between real pixels and detected is minimized [20].

Canny detector is composed of 4 steps:

1) *Gaussian filter*: Is applied to reduce noise and get and smooth image by convolution with the kernel matrix in equation (9).

$$K = \frac{1}{159} \begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix} \quad (9)$$

2) *Intensity gradient*: Is obtained calculating  $\left|\frac{\partial f}{\partial x}\right|$  and  $\left|\frac{\partial f}{\partial y}\right|$  using equations (7) and (8) and substituting in equation (4). Then gradient angle is calculated with equation (10).

$$\theta_G = \arctan \left( \frac{\left|\frac{\partial f}{\partial y}\right|}{\left|\frac{\partial f}{\partial x}\right|} \right) \quad (10)$$

3) *Comparison*: The pixels obtained after perform the gradient are compared with discrete angle  $\theta_D$  which can be  $0^\circ, 45^\circ, 90^\circ$  or  $135^\circ$ , depending of the pixel neighbourhood, the  $\theta_D$  is selected and if the rounded  $\theta_G$  satisfies  $round(\theta_G) > \theta_D$ , then the pixel belongs to the edge.

4) *Threshold*: Pixels are evaluated with two thresholds to avoid hysteresis or the duplicity of data. Pixel is accepted if gradient is greater than upper threshold. Pixel is rejected if gradient is below the lower threshold. If gradient is between upper and lower thresholds is accepted only if is connected to a pixel with gradient over the upper threshold.

### D. Erode and Dilate Operations

Erode and dilate operations are the most common morphological operations in image processing [20].

Erosion operation makes a 2D convolution with a binary picture and decrease the white content pixels with value of 1 depending of the size of the kernel. If kernel is greater, then more pixels are changed to black or 0 value (Fig. 7).

Dilate operation makes the opposite to erosion, which means that zero values near to white pixels are changed to white or to value of 1, depending of the used kernel (Fig 7).

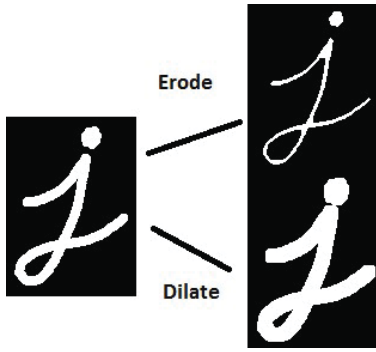


Fig. 7. Erode and Dilate Operations.

## V. METHODOLOGY

### A. Controller

Feedback controllers are commonly known by its main advantage of perturbation resistance [21]. Nao robot also requires a feedback controller, since it is not possible to make it walk following a desired path without sense its position, because of the external perturbations like floor friction and irregularities in the path.

When a robot is walking following a path, there are several lines parallel to the followed path, because every object and person will avoid to block it like is shown in [22].

Since all lines in left side of the path are parallel to the left path line and all lines in the right are parallel to the right path line. Left and right lines are equals in the slope magnitude but with opposite direction, when the robot is placed in the middle of the path lines, like is shown in Fig. 8.

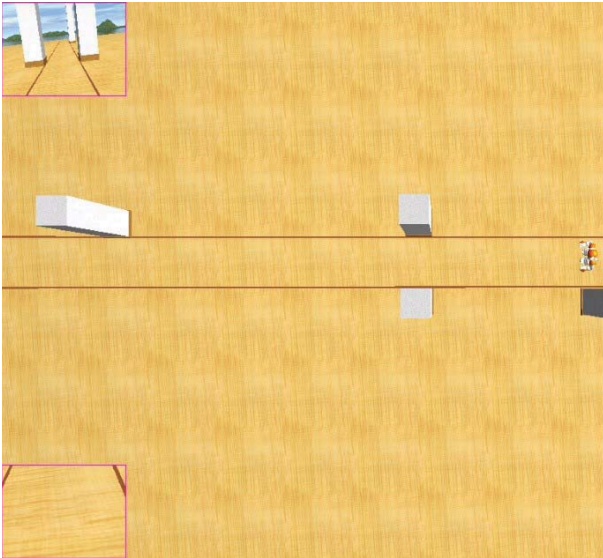


Fig. 8. Nao robot placed in the middle of the path.

Which means that when Nao is placed in the center of the path and is following it with the correct orientation, the equation (11) will be fulfilled.

$$m_1 + m_2 = 0 \quad (11)$$

With  $m_1$  and  $m_2$  representing the slope of the detected lines of the followed path, with the equation of the line  $i$  represented by equation (12).

$$y_i = m_i x_i + b_i \quad (12)$$

If the coordinates of the points that describe the line  $i$  are known, then  $m$  of every slope can be obtained with equation (13).

$$m_i = \frac{y_{i2} - y_{i1}}{x_{i2} - x_{i1}} \quad (13)$$

Then, when the slopes of the lines in the path to follow are known, will be possible to follow the path by proposing a controller that make  $m_1 + m_2$  get values near to zero.

Our proposed controller modifies the direction where Nao walks according to the  $m_1 + m_2$  sum. When  $|m_1| > |m_2|$  the robot is oriented to path line 1 and if  $|m_1| < |m_2|$  is oriented to path line 2.

Considering that the walk function for Nao requires the distance to walk in meters for “x” and “y” axes, and the rotation angle around “z” represented with “ $\theta_R$ ” in radians, according to its reference coordinate system, where the “x” axe points to Nao front side, “y” axe points to left hand and “z” axe points upwards, like is shown in Fig 9. Then the direction where the robot must walk  $\theta_R$  can be changed according to equation (14).

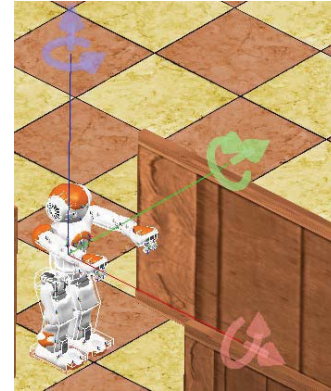


Fig. 9. Nao coordinate system.

$$\theta_R = (m_1 + m_2)P \quad (14)$$

With  $P$  as a proportional value adjusted according to the requirements of the robot.

Then our proposed structure for a walking direction controller is shown in Fig. 10. The acquired pictures from camera of Nao are processed with de algorithm described in section 5.2, which obtains the slope 1 ( $m_1$ ) and slope 2 ( $m_2$ ) required to apply equation (14), finally the direction angle obtained ( $\theta_R$ ) is used to make walk the Nao robot, with a constant distance in “x” axis.

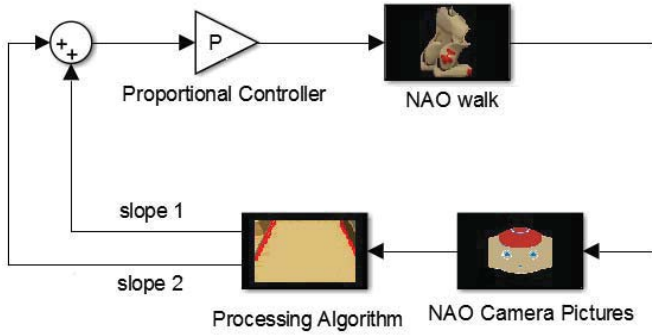


Fig. 10. Nao walking direction controller.

### B. Processing Algorithm

The objective of the processing algorithm like was mentioned before is to detect the slopes of the path lines located in the picture.

Detecting the edges in the picture obtained by Nao is the first step so that lines could be detected, but this process must acquire only vertical lines and ignore horizontal lines, because this horizontal lines will be in the departure and finish lines (Fig. 2).

Sobel Operator with vertical lines must be implemented for detecting intensity variations in  $y$  following the process in section 4.2, allowing ignoring horizontal lines.

After apply Sobel operator then Canny detecting edge filter is applied with the procedure in section 4.3, this will obtain the vertical edges detected with Sobel and will reduce noise and duplicity of data.

Erode and Dilate operations in section 4.4 are applied with a  $2 \times 2$  kernel of ones to delete remains of noise and increase the size of important data, then Sobel and Canny operators are applied again, obtaining clear vertical lines without noise and duplicity.

Then Hough line transform with process in 4.1 is applied to obtain the coordinates of start and end,  $(x_{i1}, y_{i1})$  and  $(x_{i2}, y_{i2})$  respectively of the  $i$  detected lines.

After obtaining the lines components, all the slopes  $m_i$  are obtained using equation (13) and separated in positive slopes and negative slopes, so that be possible to find which slopes belongs to the right line path and the left line path.

Finally,  $m_1$  is obtained with the average of the positive slopes and  $m_2$  with the average of negatives using equations (15) and (16).

$$m_1 = \frac{\sum_{k=1}^a m_{k-left}}{a} \quad (15)$$

$$m_2 = \frac{\sum_{k=1}^b m_{k-right}}{b} \quad (16)$$

Where  $a$  is the number of positive slopes,  $b$  the one of negative slopes,  $m_{k-left}$  are the left path line slopes and  $m_{k-right}$  the right path line slopes.

### C. General Procedure

In general, the procedure steps to follow the desired path in the run event of the Nao are:

1) *Implement acquisition of images:* Obtained from the low camera of Nao, using python 2.7 and naoqi 1.14.5 library, which is used to link the PC with the robot.

2) *Apply the processing algorithm (described in 5.2):* With the obtained images, where all filters and required operations are implemented in python using Opencv 2.4.9 library for computer vision.

3) *Calculate Direction angle:* Use the obtained slopes ( $m_1$  and  $m_2$ ) in previous step to calculate the direction angle  $\theta_R$  using equation (14).

4) *Walk Parameters:* Use walk function from naoqi 1.14.5 to make walk Nao Robot with a distance in “y”  $d_y$  of 0 meters, a constant distance in “x”  $d_x$  adjusted to get the good results, following the  $\theta_R$  direction angle obtained in step 4.

The Fig. 11 shows a flow diagram of the algorithm implemented.

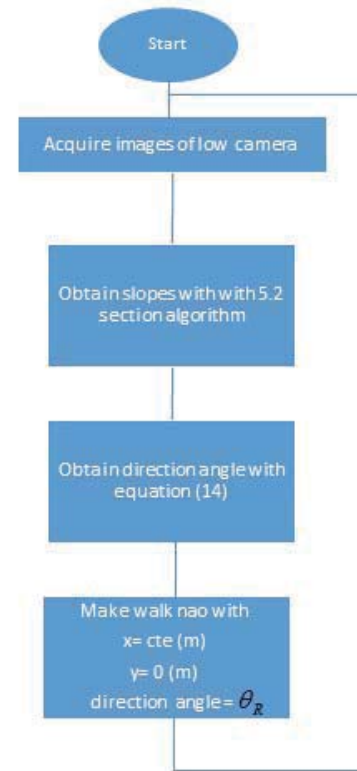


Fig. 11. Flow diagram of the path follower algorithm proposed.

VI. RESULTS

Proposed algorithm was simulated using nao webots 8.2, where a wood floor is applied in order to use a similar texture to the implemented in the contest.

The lines detected with the processing algorithm described in B of section V are shown in Fig. 12.

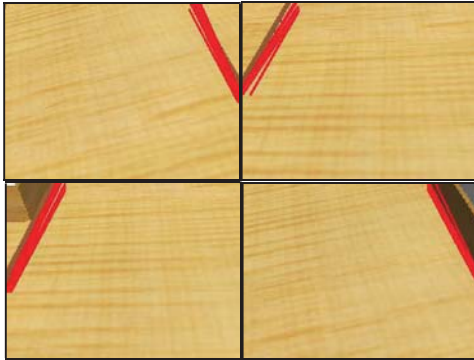


Fig. 12. Lines detected with processing algorithm in section V.

Since Nao doesn't have any GPS or other sensor that measure directly the robot position, it is not possible to implement an algorithm that can verify the real or simulated response of the controller.

In order to solve this problem, the top view of simulation was recorded without move the camera view point, so that the only object with movement was the robot, making possible to subtract from the picture when the robot is not visible, every frame of the video. All frames were changed from red, green and blue space unsinged integer 8 bits (RGB uint8) to grey scales, because colours are not required to get the position of the robot with the subtracted pictures.

The result of the subtraction creates a new image which is changed to binary space by rounding pixels divided by  $2^8 - 1 = 255$  corresponding to the 8 bits used to represent the pixels.

Then a dilate operation is applied and finally centroid algorithm is used to get the location of the robot in the video.

The picture where the robot is not visible is plotted with a red point in the coordinates of every detected centroid.

Finally, the position of the robot is obtained by multiplying the centroid coordinates  $(x_c, y_c)$  per the relation between the known sizes of the path in meters  $(s)$  and pixels  $(s_p)$ , in other words the robot position coordinates  $(x, y)$  in meters are

$$\text{obtained with } x = x_c \left( \frac{s}{s_p} \right) \text{ and } y = y_c \left( \frac{s}{s_p} \right).$$

Matlab™ 2013 was used to process the simulation video performing the required image operations.

The results showing the detection of the robot and the deviation in meters (m) of the original position (horizontal walked distance and deviathion in the path), while processing the simulation video are shown from Fig. 13 to 16.

Speed and average speed of the robot are calculated with the obtained position and its change in the time.

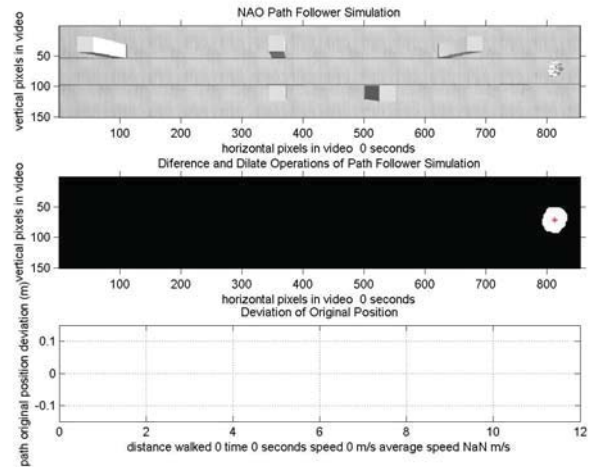


Fig. 13. Robot located in the video simulation at 0 seconds.

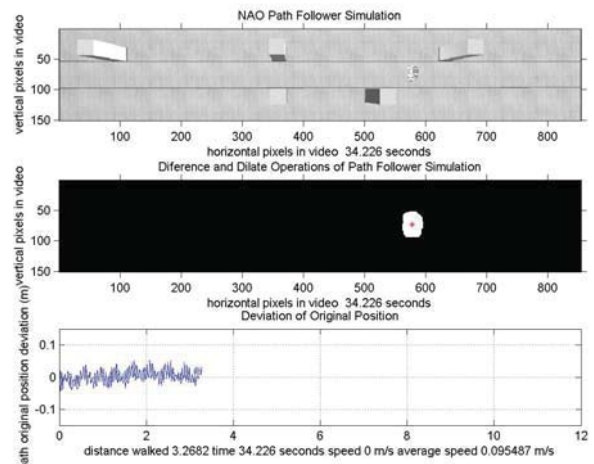


Fig. 14. Robot located in the video simulation at 34.226 seconds.

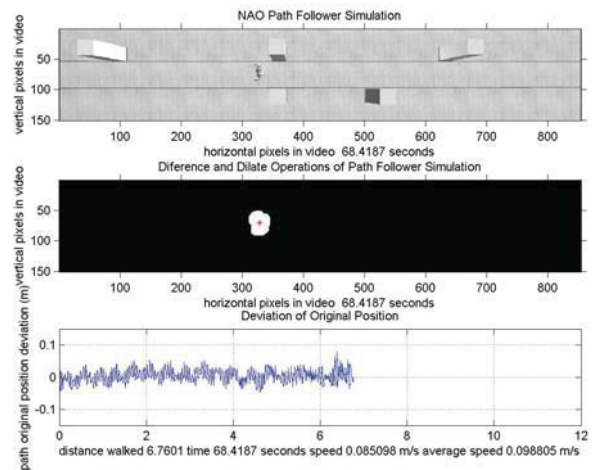


Fig. 15. Robot located in the video simulation at 68.4187 seconds

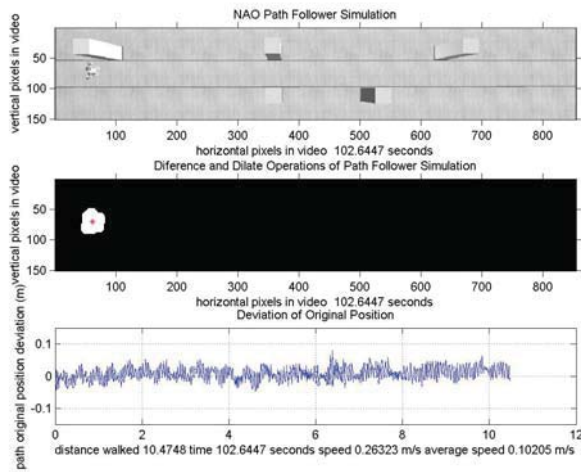


Fig. 16. Robot located in the video simulation at 102.6447 seconds

The trajectory followed by the robot after simulation using our proposed algorithm is shown in Fig 17.

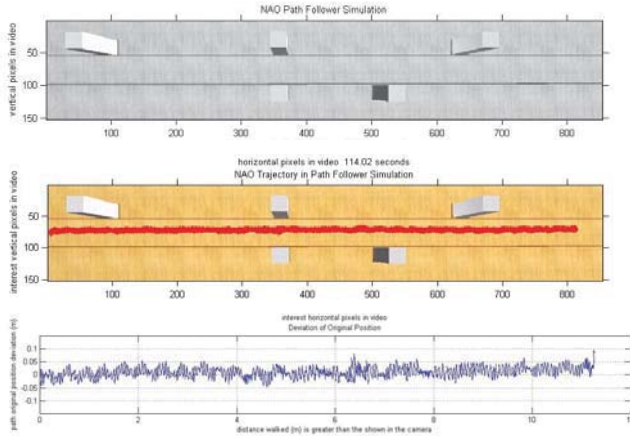


Fig. 17. Robot trajectory followed with the proposed algorithm.

The real implementation pictures obtained with the processing algorithm described in section V ignoring horizontal lines are shown in Fig 18 and the Nao robot following the tested path, in a tile floor of 0.5 m edge length is shown in Fig 19.

Plotting images and implementing of our processing algorithm without considering the walk time required, takes to Nao 0.10 seconds per frame.



Fig. 18. Pictures obtained of Nao robot low camera with the processing algorithm in section V.

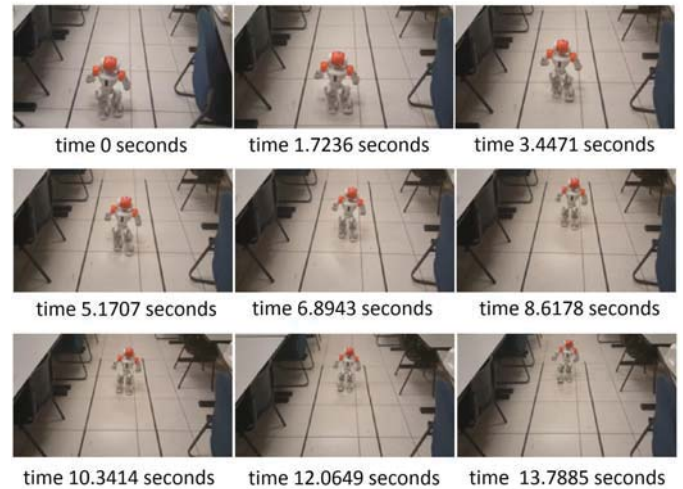


Fig. 19. Pictures of Nao following a path with tiles floor of 0.5 m edge length, using the proposed algorithm.

## VII. CONCLUSIONS

After real and simulated implementations our proposed algorithm successfully follow the common straight path in run event according to the specs in Robotics and Artificial Intelligence Nao 2015 contest.

Our proposed algorithm used with a proportional controller follows the desire trajectory without touching the path lines like it is shown in Fig. 17 and Fig. 19, even when Nao robot displacement using walk function is not accurate.

### A. Future Work

Development of a run algorithm without use the Nao default function for walking, resulting in the improvement of the robot speed, required to better results in the contest.

Oscillations reduction and a better path following saving time in the Nao displacement could be reached by testing PID and fuzzy controllers.

## REFERENCES

- [1] A. Barrientos, L. F. Peñín, C. Balaguer and R. Aracil, *Fundamentos de Robótica*, Madrid, España: McGraw-Hill, 1997.
- [2] M. Montes Rivera and M. O. Aguilar Justo, "Maze Solver Algorithm: for a Nao Humanoid Robot," in *Nuevos Avances en Robótica y Computación*, La Paz, Baja California Sur, México, INSTITUTO TECNOLÓGICO DE LA PAZ, 2015, pp. 268-271.
- [3] B. Coppin, *Artificial Intelligence Illuminated*, Sudbury, Massachusetts: JONES AND BARTLETT PUBLISHERS, 2004.
- [4] E. N. Gama Melo, O. F. Avilés Sánchez and D. Amaya Hurtado, "Anthropomorphic robotic hands: a review," *Ingeniería y Desarrollo*, vol. 32, no. 2, pp. 279-313, 2014.
- [5] V. Santos, R. Moreira and F. Silva, "Mechatronic Design of a New Humanoid Robot with Hybrid Parallel Actuation," *International Journal of Advanced Robotic Systems*, vol. 9, no. 119, 2012.
- [6] Nao Mediatec, *Manual de Jueces Concurso Nao 2014 de Robótica e Inteligencia Artificial*, Aguascalientes: Nao Mediatec, 2014.
- [7] Aldebaran Robotics, "Nao Datasheet," 06 1 2015. [Online]. Available: [https://www.aldebaran.com/sites/aldebaran/files/nao\\_datasheet.pdf](https://www.aldebaran.com/sites/aldebaran/files/nao_datasheet.pdf).

[Accessed 06 01 2015].

- [8] Mediatec, “Concurso de Robótica e Inteligencia Artificial NAO 2015,” in *Bases del concurso*, Aguascalientes, 2015.
- [9] A. M. Zhang and R. A. Russell, “Dominant Orientation Tracking for Path Following,” *Internacional Conference on Intelligent Robots and Systems*, pp. 3885-3889, 2005.
- [10] M. Pakdaman, M. M. Sanaatiyan and M. R. Ghahroudi, “A Line Follower Robot from design to Implementation: Technical issues and problems,” *The 2nd International Conference on Computer and Automation Engineering (ICCAE)*, pp. 5-9, 2010.
- [11] M. Engin and D. Engin, “PATH PLANNING OF LINE FOLLOWER ROBOT,” *Proceedings of the 5th European DSP Education and Research Conference*, pp. 1-5, 2012.
- [12] A. Fujimori, T. Fujimoto and G. Bohatscs, “Distributed Leader-Follower Navigation of Mobile robots,” *International Conference on Control and Automation*, vol. 2, pp. 960-965, 2005.
- [13] A. Loria, J. Dardemir and N. A. Jarquin, “Leader-Follower Formation and Tracking Control of Mobile Robots Along Straight Paths,” *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY*, p. 1, 2015.
- [14] M. Burke and W. Brink, “Estimating Target Orientation with a Single Camera for Use in a Human-Following Robot,” *Proceedings of the Twenty-First Annual Symposium of the Pattern Recognition Association of South Africa*, 2010.
- [15] M.-S. Lee, M.-J. Jung and J.-H. Kim, “Evolutionary Programming-based Fuzzy Logic Path Planner and Follower for Mobile Robots,” *Proceedings of the 2000 Congress on Evolutionary Computation*, vol. 1, pp. 139-144, 2000.
- [16] G. Cu, A. G. Ang, A. R. Ballesteros and J. C. Rentoy, “HUMAN FOLLOWING ROBOT USING KINECT SENSOR,” *Research Congress 2013*, 2013.
- [17] J. Delfin, H. M. Becerra and G. Arechavaleta, “Visual Path Following using a Sequence of Target Images and Smooth Robot Velocities for Humanoid Navigation,” *International Conference on Humanoid Robots (Humanoids)*, pp. 354-359, 2014.
- [18] P. Hough, “Method and Means for Recognizing Complex Patterns”. Estados Unidos de America Patent 3069654, 1962.
- [19] D. Antolovic, “INDIANA UNIVERSITY BLOOMINGTON,” 15 04 2008. [Online]. Available: <http://www.cs.indiana.edu/pub/techreports/TR663.pdf>. [Accessed 08 02 2016].
- [20] A. Bovik, *Hand Book of Image and Video Processing*, San Diego: Academic Press, 2000.
- [21] K. Ogata, *Ingeniería de Control Moderna*, Madrid: PEARSON EDUCACIÓN, 2010.
- [22] D. Nair and J. K. Aggarwal, “Moving Obstacle Detection From a Navigating Robot,” *IEEE TRANSACTIONS ON ROBOTICS AND AUTOMATION*, vol. 14, pp. 404-416, 1998.

# Sobre el Cálculo del ZMP en la Estrategia Tobillo para el Balance de un Robot Humanoide

Bahena-García L., Bugarin E. y Aguilar-Bustos A. Y.

Tecnológico Nacional de México. Instituto Tecnológico de Ensenada  
Ensenada, Baja California, México

Correo-e: {lbahena, ebugarin, aguilar}@ite.edu.mx

**Resumen**—En este documento se presenta un análisis relacionado con la estrategia Tobillo para el balance dinámico de postura de un robot humanoide sobre el “Punto de Momento Cero (ZMP por sus siglas en inglés) calculado”. Se diseña un prototipo que consiste en un sistema péndulo invertido para emular un pie y una pierna (y cuerpo del robot), en donde, a manera de tutorial, su modelo dinámico se obtiene utilizando el método Newton-Euler. El propósito fundamental de este trabajo es mostrar experimentalmente, sin medición directa de las fuerzas de reacción del piso o suelo sobre el pie, que mientras el “ZMP calculado” se encuentre dentro del polígono de soporte (en este caso de estrategia Tobillo, dentro de la superficie de contacto del pie con el piso o suelo) no habrá rotación del pie respecto a sus bordes, lo que a final de cuentas podría representar un desequilibrio en la postura erguida del robot humanoide.

**Palabras clave**—Balance dinámico de postura; Estrategia Tobillo; ZMP calculado; Robots humanoides; Validación experimental.

## I. INTRODUCCIÓN

La locomoción bípeda es uno de los problemas más interesantes a resolver dentro del campo de los robots humanoides, la cual consta de dos fases: la de soporte único y la de soporte doble [1]. La fase de soporte único se tiene cuando un solo pie está en contacto con el piso o suelo y la fase de soporte doble es cuando los dos pies están en contacto con el suelo. Un par o una fuerza no balanceada en el robot puede llevar al desequilibrio de su postura erguida y eventualmente a una posible caída. Para resolver este problema, infinidad de trabajos han utilizado el criterio del Punto de Momento Cero (ZMP por sus siglas en inglés), el cual fue introducido por Vukobratovic en 1972 (ver [2]). Básicamente, el ZMP es una medida para juzgar si el contacto entre el pie (o los pies) del robot y el suelo puede mantenerse y se define como el punto en el suelo donde la resultante de las fuerzas de reacción del suelo  $R$  actúa (ver Fig. 1).

Un concepto muy importante en este esquema es el de polígono de soporte, ya que el ZMP solo existe dentro de éste. El polígono de soporte se define como la región que se formaría si se encerraran todos los puntos de contacto del robot con el suelo mediante un cordón elástico [3]. En el caso de la fase de soporte único el polígono de soporte es el área de contacto del pie del robot con el piso o suelo.

Otro criterio que puede utilizarse como herramienta para evaluar el balance de postura en robots con patas, y específicamente en robots humanoides, es el punto Indicador

de Rotación del Pie (FRI por sus siglas en inglés), el cual se define como el punto en la superficie de contacto entre el pie y el suelo, dentro o fuera del polígono de soporte, en donde el momento resultante de la fuerza/par aplicado al pie es normal a la superficie [4]. El FRI sí puede abandonar el polígono de soporte, contrario al ZMP que sólo existe dentro de éste. Si el punto FRI se encuentra dentro del polígono de soporte entonces es igual al ZMP. Cuando el punto FRI se encuentra fuera del polígono de soporte habrá rotación alrededor de los bordes del pie y entre más se aleje de los límites del pie, más grande será la rotación y mayor el desbalance.

Por otra parte, el “ZMP calculado” se define de la misma manera que el ZMP pero considerando que el pie o los pies se encuentran pegados o atados al suelo, de esta forma el “ZMP calculado” sí pudiera encontrarse fuera del polígono de soporte [3].

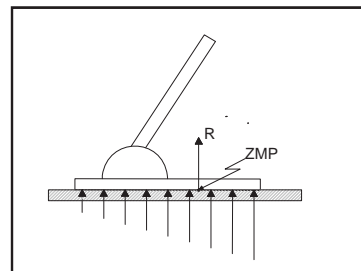


Fig. 1. Definición del ZMP.

Ahora bien, en el balance de postura erguida de un robot humanoide pueden distinguirse, básicamente, tres estrategias [5-9]: la de Tobillo, la de Cintura y la de Paso Reactivo. Como sus propios nombres lo indican, en la primera y segunda estrategia todo el robot debe considerarse rígido y sólo se guarda su postura erguida mediante movimientos en la articulación del tobillo o de la cintura, respectivamente; mientras que en la tercera estrategia el robot da un paso para equilibrarse. Como es de notarse, en la Estrategia Paso Reactivo se soportan mayores fuerzas o pares desequilibrantes, en tanto que la Estrategia Tobillo es la que soporta una menor cantidad de dichas fuerzas o pares. Estudios en Biomecánica y Rehabilitación muestran que estas tres estrategias son las que efectivamente el cuerpo humano utiliza para mantener su postura erguida.

En el presente trabajo se describe un análisis, validado experimentalmente, sobre el “ZMP calculado” en la Estrategia Tobillo. Para tal efecto se diseña un prototipo (ver Fig. 2) que consiste en un sistema péndulo invertido para emular un pie y una pierna (y cuerpo del robot). A manera de tutorial, su modelo dinámico se obtiene utilizando el método Newton-Euler tal y como se explica en [10]. De esta manera, el propósito fundamental de este trabajo es mostrar experimentalmente, sin medición directa de las fuerzas de reacción del piso o suelo sobre el pie, que mientras el “ZMP calculado” se encuentre dentro del polígono de soporte (en este caso de estrategia Tobillo, dentro de la superficie de contacto del pie con el piso o suelo) no habrá rotación del pie respecto a sus bordes, lo que a final de cuentas podría representar un desequilibrio en la postura erguida del robot humanoide. En el análisis se supondrá que la fricción entre el pie y el suelo es suficiente para que no haya deslizamiento.

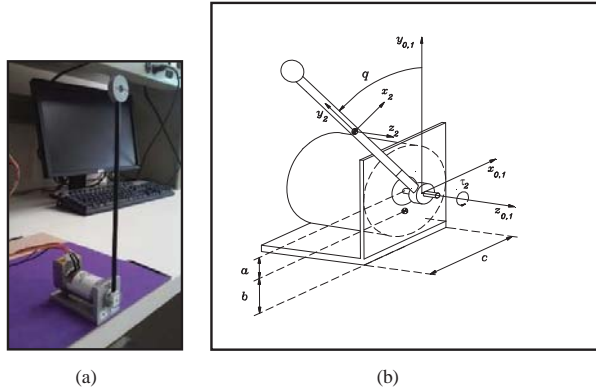


Fig. 2. Sistema utilizado (a) fotografía y (b) esquema.

## II. MODELO DINÁMICO

Como puede observarse en la Fig. 2, el sistema está compuesto por un motor de corriente directa y un péndulo atado a su eje. Consecuentemente, el sistema puede separarse en dos cuerpos libres (ver figuras 3 y 4): (a) el estator y la base (el pie) y (b) el rotor y el péndulo (la pierna y el cuerpo del robot). Sus parámetros y variables se describen en la Tabla I.

El esquema de la Fig. 2(b) muestra cómo fueron colocados diferentes marcos coordenados al sistema:  $\Sigma_0$  es el marco inercial,  $\Sigma_1$  es el marco del pie y  $\Sigma_2$  es el marco de la pierna y el cuerpo del robot. Las matrices de rotación correspondientes entre dichos marcos son:

$$R_1^0 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ y } R_2^1 = \begin{bmatrix} \cos(q) & -\sin(q) & 0 \\ \sin(q) & \cos(q) & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

### A. Subsistema mecánico

A manera de tutorial, en este apartado se obtendrá el modelo dinámico del subsistema mecánico recién descrito siguiendo el método recursivo Newton-Euler, tal y como se presenta en [10]. De esta forma, se tienen dos cuerpos rígidos, por lo que  $n = 2$ .

Ahora, para encontrar las velocidades y aceleraciones de cada cuerpo rígido del sistema se resuelve recursivamente hacia enfrente, en este caso para  $i = 1, 2$ . Primero se parte de las condiciones iniciales

$$\omega_0 = 0, \alpha_0 = 0, a_{c,0} = 0 \text{ y } a_{e,0} = 0$$

donde  $\omega_i$  y  $\alpha_i$  son la velocidad y aceleración angular del cuerpo  $i$  expresadas en  $\Sigma_0$ , respectivamente; y  $a_{c,i}$  y  $a_{e,i}$  son la aceleración del centro de masas y del extremo del cuerpo  $i$  expresadas en  $\Sigma_i$ , respectivamente.

Para  $i = 1$ :

$$\omega_1 = 0, \alpha_1 = 0, a_{c,1} = 0, a_{e,1} = 0;$$

es decir, se considera que el pie no tiene movimiento.

TABLA I. PARÁMETROS Y VARIABLES DEL SISTEMA.

Parámetro	Descripción
$m_1$	Masa del pie (estator y base)
$m_2$	Masa de la pierna y cuerpo del robot (rotor y péndulo)
$g$	Aceleración debida a la gravedad
$q$	Posición angular de la pierna y cuerpo del robot
$\tau_2$	Par generado por el motor
$l$	Distancia (a lo largo de $y_2$ ) al centro de masas de la pierna y cuerpo del robot desde el origen de $\Sigma_1$
$a$	Distancia vertical al centro de masas del pie desde el origen de $\Sigma_1$
$b$	Distancia vertical entre el centro de masas del pie y el contanto del pie con el suelo
$c$	Ancho de la base o pie

Y para  $i = 2$ , claramente

$$\omega_2 = [0 \quad 0 \quad \dot{q}]^T \text{ y } \alpha_2 = [0 \quad 0 \quad \ddot{q}]^T;$$

en tanto que (ver [10])

$$a_{c,2} = (R_2^1)^T a_{e,1} + \dot{\omega}_2 \times r_{2,c2} + \omega_2 \times (\omega_2 \times r_{2,c2})$$

donde  $r_{2,c2} = [0 \quad l \quad 0]^T$  es el vector de posición desde el eje de giro del cuerpo 2 (o cuerpo  $i$  en el caso genérico, ver figuras 3 y 4) a su centro de masas (por supuesto, en la ecuación  $\dot{\omega} = \alpha$ ). Por lo que después de simplificar

$$a_{c,2} = \begin{bmatrix} -l\ddot{q} \\ -l\dot{q}^2 \end{bmatrix}.$$

Para este caso ya no es necesario encontrar  $a_{e,2}$ .

Luego, con la recursión hacia atrás, desde  $i = 2$  a  $i = 1$ , se obtienen las fuerzas  $f_i$  y pares  $\tau_i$  del sistema partiendo de las condiciones terminales

$$f_{n+1} = 0 \text{ y } \tau_{n+1} = 0.$$

Para  $i = 2$  (ver Fig. 4 y [10]):

$$f_2 = \begin{bmatrix} H_2 \\ V_2 \\ 0 \end{bmatrix} = R_3^2 f_3 + m_2 a_{c,2} - m_2 g_2 \quad (1)$$

donde  $g_i = R_0^i g_0$  es el vector de la aceleración de la gravedad expresada en el marco  $i$ . Resolviendo, tenemos que (con abuso de notación, sin considerar la tercer componente)



$$f_2 = \begin{bmatrix} H_2 \\ V_2 \end{bmatrix} = \begin{bmatrix} m_2 g \text{sen}(q) - m_2 l \ddot{q} \\ m_2 g \text{cos}(q) - m_2 l \dot{q}^2 \end{bmatrix}. \quad (2)$$

Ahora calculamos  $\tau_2$  de la siguiente forma (ver [10]):

$$\tau_2 = R_3^2 \tau_3 - f_2 \times r_{2,c2} + (R_3^2 f_3) \times r_{3,c2} + I_2 \alpha_2 + \omega_2 \times (I_2 \omega_2)$$

donde  $I_i$  es el tensor de inercia del cuerpo  $i$  respecto a un marco paralelo al marco  $i$  cuyo origen es el centro de masas del cuerpo  $i$ . De manera que simplificando se obtiene

$$\tau_2 = (m_2 l^2 + I_{zz}) \ddot{q} - m_2 g l \text{sen}(q), \quad (3)$$

lo cual es expresado respecto al eje de giro (con abuso de notación) y en donde  $I_{zz}$  es el momento de inercia respecto al eje de giro del cuerpo 2.

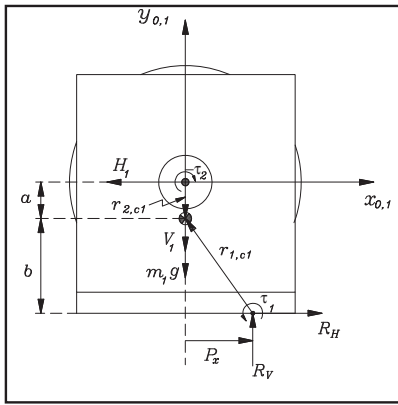


Fig. 3. Vista frontal del diagrama del cuerpo libre del pie.

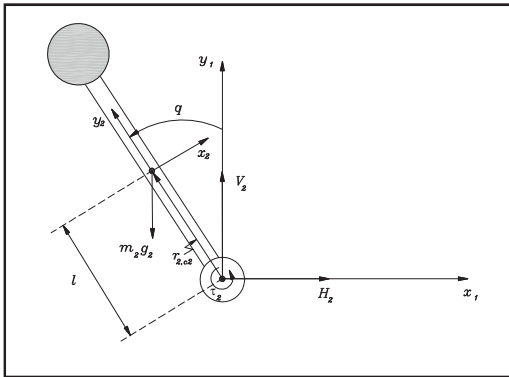


Fig. 4. Vista frontal del diagrama de cuerpo libre de la pierna y el cuerpo del robot.

Obsérvese que tomando en cuenta la generación del par  $\tau_m$  del motor de corriente directa

$$\tau_2 = \tau_m - f_m(\dot{q}), \quad (4)$$

en donde  $f_m(\dot{q})$  es la fricción del sistema.

Finalmente, para  $i = 1$  (ver Fig. 3):

$$f_1 = \begin{bmatrix} R_H \\ R_V \\ 0 \end{bmatrix} = R_2^1 f_2 + m_1 a_{c,1} - m_1 g_1. \quad (5)$$

La cual, después de simplificar (y sin escribir la tercer componente), resulta

$$f_1 = \begin{bmatrix} R_H \\ R_V \end{bmatrix} = \begin{bmatrix} m_2 l \text{sen}(q) \dot{q}^2 - m_2 l \text{cos}(q) \ddot{q} \\ (m_1 + m_2) g - m_2 l \text{sen}(q) \ddot{q} - m_2 l \text{cos}(q) \dot{q}^2 \end{bmatrix}. \quad (6)$$

El par  $\tau_1$  se calcula mediante

$$\tau_1 = R_2^1 \tau_2 - f_1 \times r_{1,c1} + (R_2^1 f_2) \times r_{2,c1} + I_1 \alpha_1 + \omega_1 \times (I_1 \omega_1), \quad (7)$$

el cual debe ser igual a cero. Por lo tanto, la tercer componente de (7) se puede simplificar a

$$0 = \tau_2 - R_H b - R_V p_x - R_H a \quad (8)$$

donde  $p_x$  es el "ZMP calculado" del sistema. Despejando  $p_x$  de (8) se obtiene que

$$p_x = \frac{\tau_2 - R_H(a+b)}{R_V}. \quad (9)$$

### B. Subsistema eléctrico

Para encontrar el par generado por el motor  $\tau_m$ , en este apartado se desarrollarán las ecuaciones del subsistema eléctrico suponiendo que la inductancia de armadura  $L_a \approx 0$ ; de esta manera, el voltaje de armadura  $v$  del motor está dado por

$$v = R_a i_a + k_b \dot{q} \quad (10)$$

donde  $i_a$  es la corriente de armadura (en la Tabla II puede revisarse la descripción de todos los parámetros eléctricos del sistema). La ecuación que relaciona el par del motor  $\tau_m$  con la corriente de armadura  $i_a$  es

$$\tau_m = k_a i_a. \quad (11)$$

De esta forma, utilizando (3), (4), (10) y (11), se puede llegar finalmente a que

$$v = \frac{R_a}{k_a} [(m_2 l^2 + I_{zz}) \ddot{q} - m_2 g l \text{sen}(q) + f_m(\dot{q})] + k_b \dot{q}. \quad (12)$$

### III. CONTROLADOR DE MOVIMIENTO

Para la medición del "ZMP calculado" se utilizará (9), de modo que no se hará uso de sensores de fuerza. Para este fin, se diseñará un controlador de movimiento para provocar en el péndulo un movimiento sinusoidal. De manera que, jugando con diferentes amplitudes y frecuencias, se pueda tener cierta certeza (al menos cerca de los bordes del pie) de la localización del "ZMP calculado", ya que se estará atento a posibles rotaciones del pie respecto a sus bordes.

Por lo tanto, defínase el siguiente objetivo de control

$$\lim_{t \rightarrow \infty} q(t) = q_d(t) \quad (13)$$

donde  $q_d(t) = A \text{sen}(2\pi f t)$  es la posición sinusoidal deseada del péndulo, con  $A$  su amplitud y  $f$  su frecuencia.

Para resolver (13) se propone la siguiente ley de control:

$$v = \frac{R_a}{k_a} [(m_2 l^2 + I_{zz})(\ddot{q}_d + k_v \dot{\tilde{q}} + k_p \tilde{q}) - m_2 g l \text{sen}(q) + k_{bv} \dot{q} + f_c \text{sign}(\dot{q})] \quad (14)$$

donde  $k_p, k_v > 0$  son la ganancia proporcional y derivativa del controlador,  $\tilde{q} = q_d - q$  es el error de posición y se ha utilizado fricción viscosa y de Coulomb en el modelo de fricción, esto es

$$f_m(\dot{q}) = f_v \dot{q} + f_c \text{sign}(\dot{q})$$

con  $f_v$  el coeficiente de fricción viscosa y  $f_c$  el coeficiente de fricción de Coulomb. El parámetro  $k_{bv} = f_v + \frac{k_a k_b}{R_a}$ .

Al sustituir (14) en (12) se obtiene la siguiente ecuación en lazo cerrado

$$\ddot{\tilde{q}} + k_v \dot{\tilde{q}} + k_p \tilde{q} = 0.$$

Por lo que se garantiza que el objetivo de control (13) es satisfecho, puesto que por diseño  $k_p, k_v > 0$ .

Debe destacarse que la aceleración angular  $\ddot{q}$  en lazo cerrado puede calcularse mediante

$$\ddot{q} = \frac{1}{m_2 l^2 + I_{zz}} \left[ \frac{k_a}{R_a} (v - k_b \dot{q}) + m_2 g l \text{sen}(q) - f_m(\dot{q}) \right] \quad (15)$$

y, a partir de (4), (10) y (11)

$$\tau_2 = \frac{k_a}{R_a} (v - k_b \dot{q}) - f_m(\dot{q}),$$

por lo que  $p_x$ , de (9), se puede calcular con sólo el conocimiento de  $q, \dot{q}$  y  $v$ .

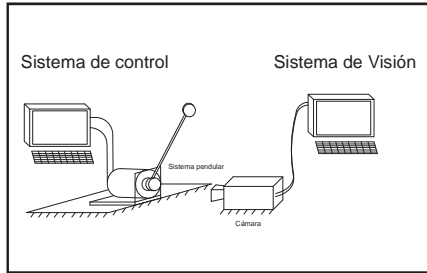


Fig. 5. Esquema de experimentación.

#### IV. SIMULACIONES Y EXPERIMENTOS

Para la experimentación se utilizó un motor de corriente directa Pittman Serie 8000 de 24 [V] puesto en una base de aluminio, según se puede observar en la Fig. 2(a). Este motor cuenta con un codificador óptico para medir su posición angular (la velocidad angular  $\dot{q}$  se estimará a partir de  $q$  mediante el método numérico de Euler). La Fig. 5 muestra el esquema de experimentación y como se puede notar se cuenta con dos sistemas de cómputo: el Sistema de Control que opera en tiempo real utilizando Matlab, bajo el sistema operativo Windows, con un periodo de muestreo de 1 [ms]; y el Sistema de Visión, el cual también trabaja en tiempo real haciendo uso de RTLinux a una cadencia de 10 [ms]. El Sistema de Visión se utiliza para medir el ángulo  $\theta$  de rotación del pie respecto a

sus bordes y de esta manera detectar que el “ZMP calculado” se encuentra fuera del polígono de soporte.

En la Tabla II se concentran los valores de los parámetros utilizados tanto en la simulación como en la experimentación. En tanto que en la la Tabla III se detallan los resultados obtenidos en los diferentes escenarios analizados. Los experimentos y simulaciones corresponden en todos los escenarios con una  $k_p = 250$ ,  $k_v = 10$ ,  $f = 2.5$  [Hz] y una  $k_{bv} = 5.6733 \times 10^{-5}$ . Como puede observarse de la Tabla III, en los primeros 3 escenarios el  $p_x$  se mantiene dentro del polígono de soporte tanto en la simulación como en la experimentación, puesto que  $p_x < c/2 = 0.0185$  [m]. Ya en el escenario 4, el  $p_x > c/2$ ; sin embargo, el sistema de visión no registra movimiento del pie ( $\theta = 0$ ), lo cual puede ser debido a que el movimiento del pie respecto a sus bordes es demasiado pequeño como para ser detectado por el Sistema de Visión utilizado. El movimiento del pie es evidente para amplitudes, en la posición sinusoidal deseada, mayores a 20 [deg].

TABLA II. VALORES DE LOS PARÁMETROS UTILIZADOS.

Parámetro	Descripción	Valor
$m_1$	Masa del pie (estator y base)	0.26 [kg]
$m_2$	Masa de la pierna y cuerpo del robot (rotor y péndulo)	0.02066 [kg]
$I_{zz}$	Momento de inercia respecto al eje de giro del cuerpo 2	$3.3256 \times 10^{-04}$ [kgm <sup>2</sup> ]
$g$	Aceleración debida a la gravedad	9.81 [m/s <sup>2</sup> ]
$l$	Distancia (a lo largo de $y_2$ ) al centro de masas de la pierna y cuerpo del robot desde el origen de $\Sigma_1$	0.12 [m]
$a$	Distancia vertical al centro de masas del pie desde el origen de $\Sigma_1$	0.005 [m]
$b$	Distancia vertical entre el centro de masas del pie y el contacto del pie con el suelo	0.02 [m]
$c$	Ancho de la base o pie	0.037 [m]
$R_a$	Resistencia de armadura	12.1 [Ω]
$k_a$	Constante motor-par	0.037 [Nm/A]
$k_b$	Constante contra-electromotriz	0.037 [Vs/rad]
$f_c$	Coefficiente de fricción de Coulomb	0.0024 [Nm]
$f_v$	Coefficiente de fricción viscosa	$1 \times 10^{-06}$ [Ns]

TABLA III. RESULTADOS

Escenario	A [°]	$p_{x_{max}}$ [m] simulado	$p_{x_{max}}$ [m] experimental	$\theta_{max}$ [°]
1	10	0.0123	0.0116	0
2	12	0.0147	0.0133	0
3	14	0.017	0.0163	0
4	16	0.0194	0.0188	0
5	18	0.0217	0.0199	0
6	20	0.0239	0.0227	0.1
7	22	0.0261	0.0242	0.15
8	24	0.0283	0.0251	0.4
9	26	0.0304	0.0273	1.3
10	28	0.0325	0.0307	3.5

Las figuras 6, 7 y 8 detallan las gráficas obtenidas de los escenarios 5, 8 y 10, respectivamente. En cada gráfica se superponen los resultados experimentales a los obtenidos en simulación, como puede observarse ambas gráficas en todos los casos son muy similares.

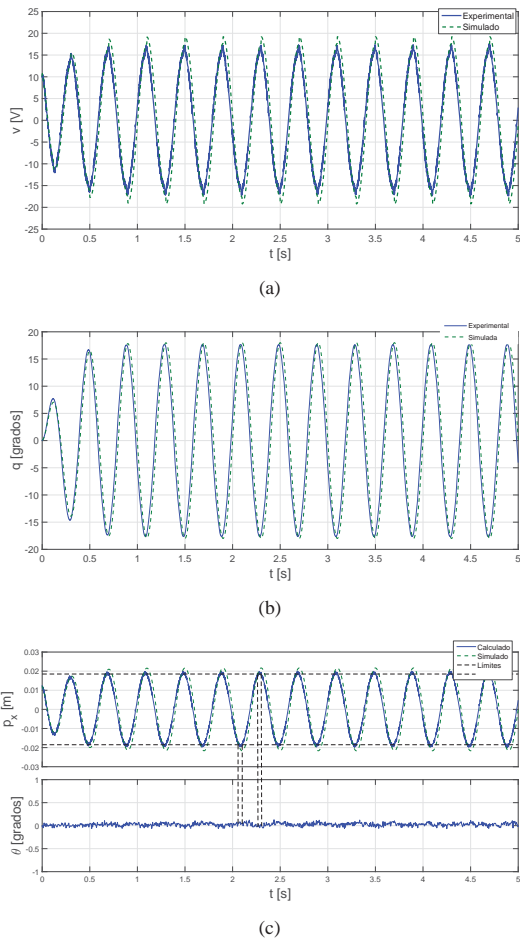


Fig. 6. Gráficas del escenario 5: (a)  $v(t)$ , (b)  $q(t)$  y (c)  $p_x(t)$  y  $\theta(t)$ .

El escenario 5 representa el límite, bajo experimentación, en donde no se registra movimiento o giro del pie respecto a sus bordes, esto es,  $\theta$  se mantiene en cero a pesar de que  $p_x$  está muy cerca de su valor máximo permitido (como se puede apreciar en la Fig. 6(c)).

En tanto que en el escenario 8, ya se ven movimientos o giros del pie, en promedio, de más de un cuarto de grado (según se presenta en la Fig. 7(c)).

Finalmente, la Fig. 8 ilustra el escenario 10 en donde el movimiento o giro del pie es muy evidente, alcanzando valores de  $\theta$  de más de 3 grados de magnitud. En este escenario se pueden notar mayores diferencias entre lo simulado y lo experimental, pero esto es debido a que en el modelado se ha supuesto que el pie se mantiene sin movimiento, lo cual, indiscutiblemente en este escenario no sucede. De hecho se registran señales no sinusoidales en  $q$  y  $v$  por esta misma razón.

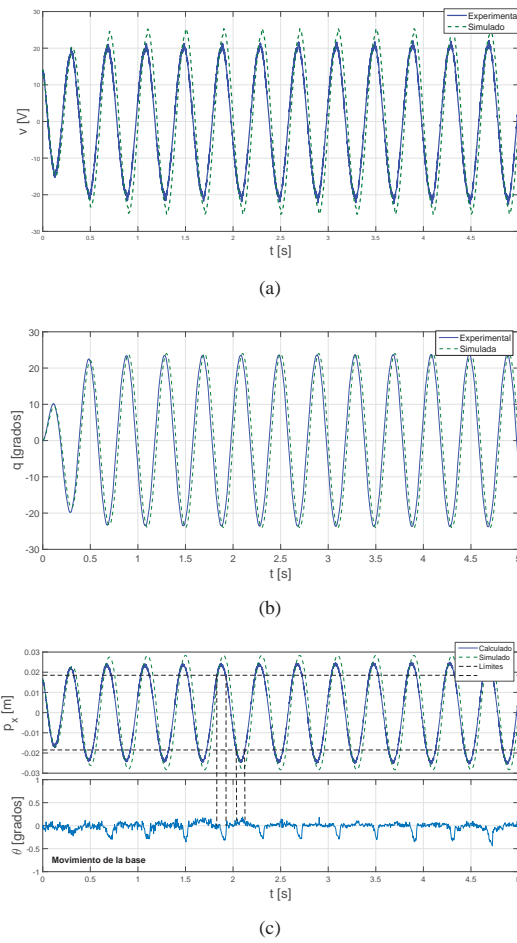


Fig. 7. Gráficas del escenario 8: (a)  $v(t)$ , (b)  $q(t)$  y (c)  $p_x(t)$  y  $\theta(t)$ .

## V. CONCLUSIONES

En este trabajo se detalla una forma de medir experimentalmente el “ZMP calculado” bajo la Estrategia Tobillo para el balance dinámico de un robot humanoide. Se ha diseñado un prototipo para emular el pie, pierna y cuerpo del robot mediante un péndulo invertido que es accionado por un motor de corriente directa. Tanto el Sistema de Control, utilizado para hacer oscilar la pierna y cuerpo del robot, como el Sistema de Visión, quien auxilia para verificar el movimiento o giro del pie respecto a sus bordes, son sistemas computacionales de tiempo real que operan a una cadencia de 1 [ms] y de 10 [ms], respectivamente. Los resultados muestran que no existe movimiento o giro del pie mientras el “ZMP calculado” esté dentro del polígono de soporte. Bajo el método de medición presentado, también se confirma que a medida que el “ZMP calculado” se aleja de los límites del polígono de soporte el pie comienza a rotar alrededor de sus bordes cada vez con una mayor intensidad, lo cual representa que el robot se encuentra en desequilibrio de su postura erguida.

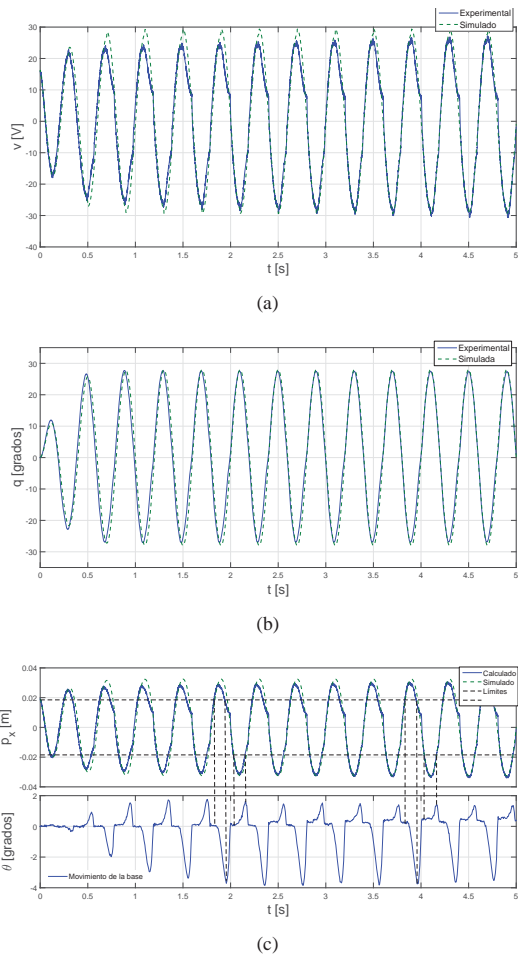


Fig. 8. Gráficas del escenario 10: (a)  $v(t)$ , (b)  $q(t)$  y (c)  $p_x(t)$  y  $\theta(t)$ .

### AGRADECIMIENTOS

Se agradece el apoyo financiero del Tecnológico Nacional de México, del Instituto Tecnológico de Ensenada, del PRODEP y del CONACYT (Proyecto 166654).

### REFERENCIAS

- [1] Westervelt, E. R., Grizzle, J. W., Chevallereau, C., Choi, J. H., and Morris, B. "Feedback control of dynamic bipedal robot locomotion". CRC Press. Taylor & Francis Group. 2007.
- [2] Vukobratovic, M. and B. Borovac. "Zero-Moment point—Thirty five years of its life". International Journal of Humanoid Robotics. Vol. 1, No.1. 2004.
- [3] Kajita, S., Hirukawa, H., Harada, K. and Yokoi, K. "Introduction to Humanoid Robotics". Springer. 2014.
- [4] Goswami, A. "Foot rotation indicator (FRI) point: A new gait planning tool to evaluate postural stability of biped robots". Proceedings of the 1999 IEEE International Conference on Robotics & Automation. Detroit, Michigan. May 1999.
- [5] Nenchev, D. N. and A. Nishio. "Ankle and hip strategies for balance recovery of a biped subjected to an impact". Robotica. Vol. 26, pp. 643—653. 2008.
- [6] Ono H., T. Sato and K. Ohnishi. "Balance recovery of ankle strategy: using knee joint for biped robot". 2011 1<sup>st</sup> International Symposium on Access Spaces (ISAS). 2011.
- [7] Kanamiya Y., S. Ota and D. Sato. "Ankle and hip balance control strategies with transitions". 2010 IEEE International Conference on Robotics and Automation. Anchorage, Alaska, USA. May 2010.
- [8] Morisawa M., F. Kanehiro, K. Kaneko, N. Mansard, J. Sola, E. Yoshida, K. Yokoi and J. P. Laumond. "Combinig suppression of the disturbance and reactive stepping for recovering balance". The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. Taipei, Taiwan. Oct 2010.
- [9] B. Stephens. "Humanoid push recovery". 2007 7th IEEE-RAS International Conference on Humanoid Robots. 2007.
- [10] Spong, M. W., S. Hutchinson and M. Vidyasagar. "Robot modeling and control". John Wiley and Sons, Inc. 2006.

# Implementación de un algoritmo de lógica difusa en el Software Matlab para un pre-diagnóstico de fallas en vehículos Volkswagen

Jennifer Olarte Villamizar<sup>1-a</sup>, Gustavo Juárez Gracia<sup>1-b</sup>

Instituto Politécnico Nacional - IPN  
Centro de Investigación en Ciencia Aplicada y Tecnología  
Avanzada - CICATA unidad Legaria  
Ciudad de México, México  
kaje.2909@gmail.com<sup>1-a</sup>, agjuarezg@ipn.mx<sup>1-b</sup>

L. Noé Oliva Moreno<sup>2</sup>

Instituto Politécnico Nacional - IPN  
Unidad Profesional Interdisciplinaria de Ingeniería  
Campus Hidalgo - UPIIH  
Ciudad de México, México  
loliva@ipn.mx<sup>2</sup>

**Resumen** —Este artículo describe el desarrollo de un algoritmo de lógica difusa utilizando el método de Mamdani y la técnica del Centroides para realizar el pre-diagnóstico de fallas de los siguientes sensores: sensores de oxígeno antes y después del catalizador, el sensor de temperatura del líquido refrigerante y las revoluciones por minuto (rpm), de un vehículo de marca Volkswagen modelo 2008 mediante el uso del software de Matlab. Los datos de entrada se escogieron por el conocimiento del técnico debido que son las variables que presentan más fallas en los vehículos de la marca Volkswagen. Los resultados obtenidos con el método de Mamdani y la técnica del Centroides, permitieron clasificar las variables en diferentes conjuntos difusos y obtener una salida difusa con las fallas que se pueden presentar en el vehículo.

**Palabras clave**—Algoritmo, Lógica difusa, Pre-Diagnóstico, Método de Mamdani, Técnica del Centroides.

## I. INTRODUCCIÓN

El número creciente de innovaciones tecnológicas en los vehículos automotores, ha obligado a los mecánicos a aprender conceptos y técnicas especializadas, enfocadas a la reparación y al mantenimiento de los mismos. La crisis energética de los años 70 y la contaminación ambiental, fue un factor importante para que esta oleada de innovaciones tecnológicas impulsaron al gobierno, a publicar estrictas normas de emisiones de gases contaminantes de los vehículos; derivado de esto, los fabricantes empezaron a desarrollar nuevos sistemas para hacer más eficientes los motores de combustión [1]. Uno de los sistemas de mayor trascendencia en la industria automotriz fue la adición de diagnósticos a bordo (OBD) en vehículos, o en términos más sencillos, la computadora que activa la luz indicadora "CHECK ENGINE" en el vehículo. El OBD I se diseñó para monitorear sistemas específicos de fabricantes en vehículos fabricados desde 1981 hasta 1995. Después surgió el desarrollo del sistema OBD II, el cual viene instalado en todos

los automóviles y camionetas a partir de 1996 que se venden en los EE.UU. Lo más relevante del OBD II, es su capacidad de aplicación universal para todos los automóviles y camionetas de años recientes de fabricación nacional y extranjera. Este programa complejo se encuentra instalado en el sistema de la computadora principal, ubicada dentro del vehículo y fue diseñado para detectar las fallas presentadas. Se puede obtener acceso a la computadora principal a través de un puerto universal OBD II, el cual se encuentra usualmente debajo del tablero [2]. A partir de estos avances, comenzó a tomar forma un modelo que se conoce como "sistema autotrónico"; que ahora se ha extendido a todos los sistemas del vehículo, y no sólo al motor [1]. Actualmente la lógica difusa se ha implementado en varias áreas referentes al ámbito automotriz como por ejemplo: evaluar la eficiencia en el modo de conducción para la reducción de combustible [3], generar un control de velocidad para entornos urbanos congestionados, entre otros [4].

La lógica difusa se ha empleado en una gran variedad de aplicaciones, como el control de procesos complejos industriales, el diseño de dispositivos artificiales de deducción automática, la construcción de artefactos electrónicos de uso doméstico y de entretenimiento, así también, en sistemas de diagnóstico [5]. Este trabajo trata sobre el desarrollo de un algoritmo de lógica difusa simulado en el software Matlab para realizar un Pre-diagnóstico de fallas, que permita detectar daños en los sensores de vehículos de marca Volkswagen con el objetivo de facilitar el diagnóstico al usuario.

## II. LÓGICA DIFUSA

### A. Definición

La lógica difusa es una rama de la inteligencia artificial que se basa en la forma en que el cerebro maneja información inexacta. Ésta lógica utiliza expresiones que no son ni totalmente ciertas ni completamente falsas, es decir, que pueden tomar un valor cualquiera de veracidad dentro de un

conjunto de valores, limitando en un extremo a la verdad absoluta y en el otro a la falsedad total [5,6]. También es un conjunto de principios matemáticos basados en grados de membresía o pertenencia, cuya función es modelar información. Este modelado se hace con base en reglas lingüísticas que aproximan una función mediante la relación de entradas y salidas del sistema. Esta lógica presenta rangos de membresía dentro de un intervalo entre 0 y 1, a diferencia de la lógica convencional, en la que el rango se limita a dos valores: el cero o el uno, (Fig. 1) [7].

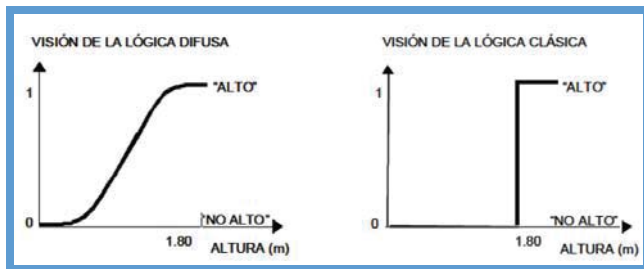


Fig.1 Lógica difusa versus lógica convencional

### B. Funciones de pertenencia

Las funciones de pertenencia pueden asignar valores en el intervalo [0,1] para definir un conjunto difuso, en el diseño de sistemas difusos suelen emplearse funciones de tipo polígonos, trapezoides o triángulos. En algunos sistemas se emplean funciones gaussianas en el centro y sigmoideas en los extremos, ya que tienen formas similares a los anteriores, y además son continuas y diferenciales, (Fig.2) [8].

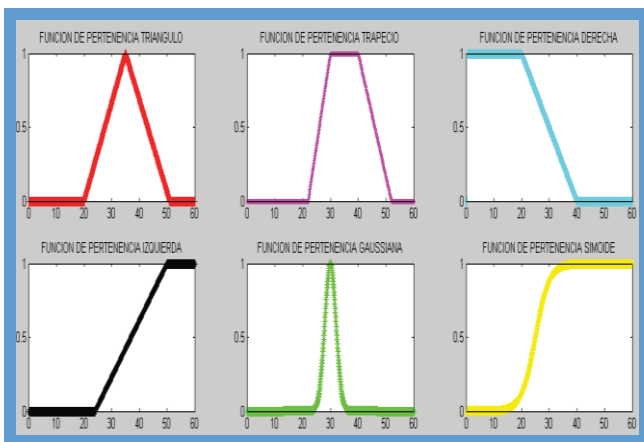


Fig.2 Algunos tipos habituales de funciones de pertenencia.

### C. Operadores lógicos difusos

Básicamente, las operaciones que se pueden realizar con conjuntos difusos, son extensiones de la lógica clásica de conjuntos, las conectivas lógicas también se definen para las operaciones de lógica difusa. Ellos están estrechamente relacionados con las definiciones de operaciones de conjuntos

difusos de Zadeh [9]. Los operadores de los conjuntos difusos de gran importancia se definen a continuación [10]:

- **Complemento:** Dado el conjunto difuso A definido en el Universo de Discurso U, su complemento es el conjunto difuso:

$$\bar{A} = \{(x, \mu_{\bar{A}}(x)) | x \in U\} \text{ y } \mu_{\bar{A}}(x) = \sim \mu_A(x) = 1 - \mu_A(x) \quad (1)$$

Donde “ $\sim$ ” denota la "negación difusa" de una función de pertenencia, coincidente en este caso con el operador de complemento para los conjuntos certeros. Hay otros operadores complemento “ $\sim$ ”. En general, se les llama C-Normas.

- **Intersección:** Dado el conjuntos difusos A y B, definidos en U con las funciones de pertenencia  $\mu_A$  y  $\mu_B$ , su intersección es el conjunto difuso:

$$A \cap B = \{(x, \mu_{A \cap B}(x)) | x \in U\}, \text{ como } \mu_{A \cap B}(x) = \mu_A(x) \sim \mu_B(x) \quad (2)$$

Siendo “ $\sim \wedge$ ” el símbolo para denotar la intersección difusa o "y difusa". Operadores de intersección pertenecen a la clase de la llamada Normas Triangulares o T-Normas.

- **Unión:** Dado los conjuntos difusos A y B, se define en U con funciones de pertenencia  $\mu_A$  y  $\mu_B$ , su unión es el conjunto difuso:

$$A \cup B = \{(x, \mu_{A \cup B}(x)) | x \in U\}, \text{ y } \mu_{A \cup B}(x) = \mu_A(x) \sim \mu_B(x) \quad (3)$$

Siendo “ $\sim \vee$ ” el símbolo para denotar la unión difusa o "o difusa". Operadores de la unión pertenecen a la clase de la llamada-S-Normas o CoNormas Triangulares (T-CoNormas).

### D. Sistemas expertos difusos

Un sistema experto difuso es un sistema basado en reglas con variables lingüísticas que emplea un mecanismo de inferencia difusa. Además de la utilización de la regla composicional de inferencia, las diferencias entre un sistema difuso y uno clásico son fundamentalmente dos [8]:

- Los sistemas difusos no suelen utilizar encadenamiento de reglas. La salida del sistema es directamente la aplicación de las reglas composicional de inferencia a las entradas.
- Las reglas del sistema se activan en “paralelo”, de modo que la salida global es una combinación de la salida de cada una de las reglas.

Otra ventaja del sistema de control basado en lógica difusa, es que no es necesario conocer un modelo matemático del sistema real, pues se puede ver como una caja negra a la cual se le proporcionan entradas, y a través del sistema esta planta generará la salida deseada. En el control convencional sí es necesario conocer la planta del sistema [11].

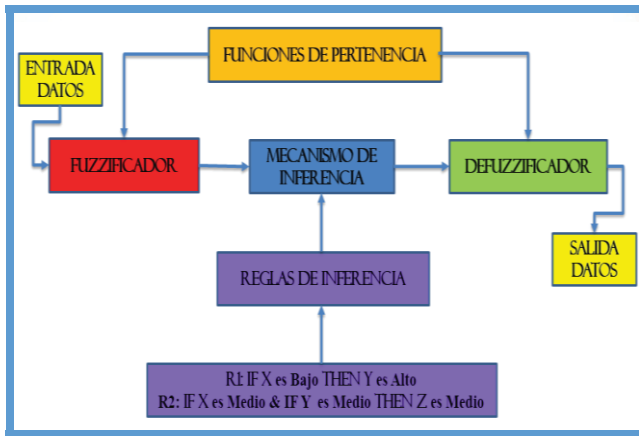


Fig.3. Elementos de un sistema difuso

En la Fig. 3 se muestra los elementos que constituyen un sistema difuso típico:

- *Fuzzificador*: El control difuso siempre involucra este proceso de Fuzzificación, esta operación se realiza en todo instante de tiempo, es la puerta de entrada al sistema de mecanismo de inferencia difusa. Es un procedimiento matemático en el que se convierte un elemento del universo de discurso (variable medida del proceso) en un valor en cada función de membresía a las cuales pertenece [12]. Así la interfaz de fuzzificación trabaja del siguiente modo:

$$A' = F(x_0) \quad (4)$$

Donde  $x_0$  es un valor preciso de entrada al sistema basado en reglas difusas, definido en el universo de discurso  $U$ ,  $A'$  es un conjunto difuso definido sobre el mismo dominio y  $F$  es un operador de fuzzificación [13].

- *Mecanismo de inferencia*: Los mecanismos de inferencia son aquellos que mediante la utilización de principios de inferencia difusa (reglas difusas), relaciona los conjuntos difusos de entrada con los conjuntos difusos de salida [7].
- *Reglas de Inferencia*: La base de reglas es la manera que tiene el sistema difuso para guardar el conocimiento lingüístico que le permite resolver el problema para el cual ha sido diseñado. Estas reglas son del tipo IF-THEN y está formada por dos partes, el antecedente y la conclusión [14].
- *Defuzzificador*: Convierte los conjuntos difusos a valores numéricos. Los resultados difusos generados por el mecanismo de inferencia no pueden ser utilizados como tal en las aplicaciones, por lo tanto, es necesario convertir las cantidades difusas en cantidades numéricas para su posterior procesamiento [15]. El proceso se denomina “defuzzificación” porque combina las variables difusas para generar el valor real de la señal (también comúnmente denominado CRISP o no-

difuso), el cual puede ser usado posteriormente para generar acciones de control [16]. A continuación se nombran los 7 diferentes métodos que existen para la obtención de los valores defuzzificados.

- (1) Principio de Max-pertenencia
- (2) Método del Centroide
- (3) Método de Promedio ponderado
- (4) Pertenencia Media-Max
- (5) Centro de Sumas
- (6) Centro de la mayor superficie
- (7) El primer máximo o la última del máximo

De los diferentes métodos de defuzzificación el método del Centroide o centro de gravedad, es el método que se va a emplear en este trabajo debido es más práctico en la implementación del hardware y es el más utilizado en el área de la lógica difusa, y se define mediante la siguiente expresión algebraica [15,17]:

$$z^* = \frac{\int_z \mu_{c'}(z)zdz}{\int_z \mu_{c'}(z)dz} \quad (5)$$

Dónde:  $\int_z \mu_{c'}(z)dz$  es el área de la conclusión difusa  $C'$  y  $z^*$  es el centro de gravedad (acción de control concreta) de la función de membresía resultante  $\mu_{c'}(z)$  de la evaluación de reglas. De manera general, la utilización de este método implica un gran esfuerzo computacional debido a la necesidad de calcular el área de la función de membresía  $\mu_{c'}(z)$  [18]. En la Fig.4 se representa gráficamente este método.

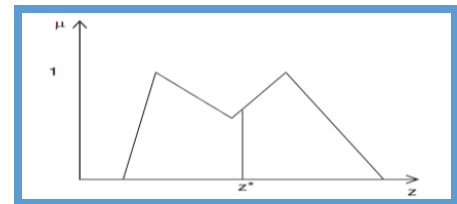


Fig.4. Método del Centroide

### III. METODOLOGÍA

Para realizar el algoritmo de lógica difusa en el software Matlab, debemos tomar en cuenta las etapas del sistema difuso de la Fig. 3. Lo primero que se realizó, fue la etapa de fuzzificación, que contiene una base de conocimiento obtenida del experto (mecánico), que proporcionó los rangos aproximados de cada una de las variables y las fallas más comunes que se presentan en estos vehículos. La base de datos se obtuvo mediante el conector OBDII, conectando un adaptador Bluetooth en la computadora del vehículo. Este adaptador se encargó de enviar los datos suministrados por el puerto de diagnóstico a una terminal de un computador, en

donde se observó la trama de datos por medio de un sniffer. Para realizar las pruebas se empleó el adaptador ELM-327 de la empresa scantool.net. Las variables con las que se trabajó fueron: el sensor de temperatura del líquido refrigerante con las rpm y los sensores de oxígeno antes y después del catalizador, las cuales presentan dos sistemas difusos expertos.

En la Fig. 5, 6, 7 y 8, se muestran los conjuntos difusos para cada una de las variables que se utilizaron en la etapa de fuzzificación, con sus respectivas funciones de pertenencia y sus etiquetas lingüísticas. En la Fig. 5 y 6 se observa el conjunto difuso para los sensores de oxígeno antes y después del catalizador, el cual se compone de tres funciones de pertenencia: pobre, normal y rico, que representan los estados de los dos sensores de oxígeno, ya que operan en los mismos rangos.

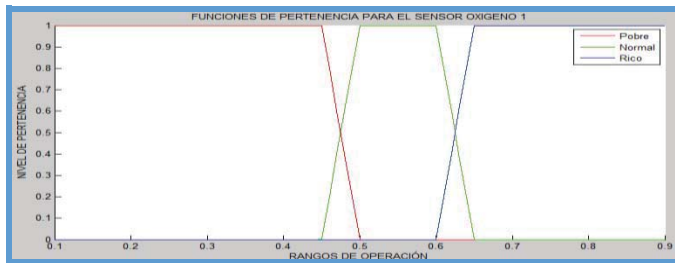


Fig.5. Conjuntos difusos para el sensor de oxígeno antes del catalizador.

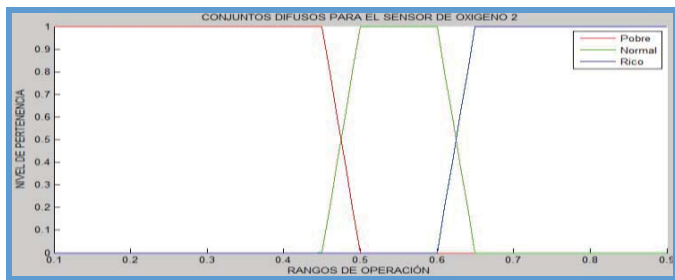


Fig.6. Conjuntos difusos para el sensor de oxígeno después del catalizador.

En la Fig. 7, se muestran los conjuntos difusos para el sensor de temperatura la cual se compone de cuatro funciones de pertenencia que representa los rangos de cada una de las etapas del sensor: frío, tibio, normal y caliente.

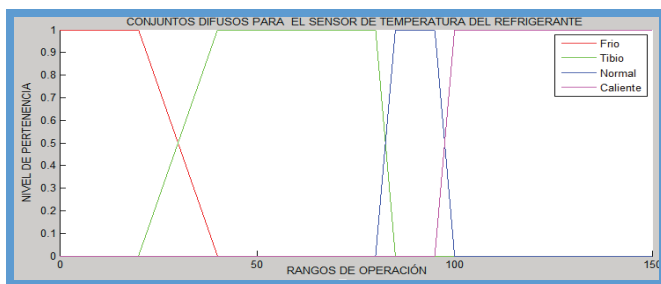


Fig.7 Conjuntos difusos para el sensor de temperatura del líquido refrigerante

En la Fig. 8, se observan los estados de las revoluciones por minuto, que en este caso se conforman del estado relenti

que significa cuando el carro esta encendido pero no está acelerado y el estado normal.

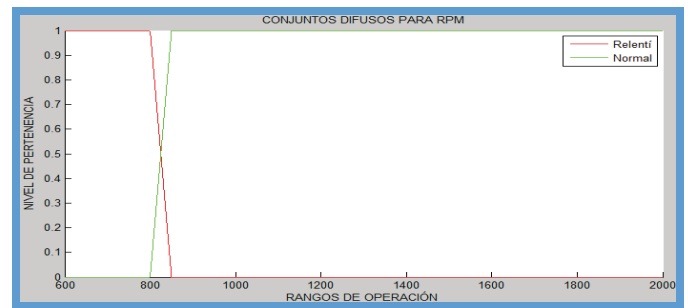


Fig.8. Conjuntos difusos para las revoluciones por minuto

En la etapa del mecanismo de inferencia se describen las reglas difusas, que asocian los conjuntos difusos de entrada versus los conjuntos difusos de salida. En la Fig. 9 y 10, se muestran el editor de las reglas difusas en el entorno de FIS (Sistema de Inferencia Difusa) que tiene Matlab, en donde las entradas son los sensores de oxígeno, antes y después del catalizador y el sensor de temperatura del líquido refrigerante con las rpm.

1. If (Sen-1 is Pobre-s1) and (Sen-2 is Probre-S2) then (output1 is Reparar) (1)
2. If (Sen-1 is Pobre-s1) and (Sen-2 is Normal-S2) then (output1 is S1) (1)
3. If (Sen-1 is Pobre-s1) and (Sen-2 is Rico-S2) then (output1 is S1) (1)
4. If (Sen-1 is Normal-S1) and (Sen-2 is Probre-S2) then (output1 is S2) (1)
5. If (Sen-1 is Normal-S1) and (Sen-2 is Normal-S2) then (output1 is Normal) (1)
6. If (Sen-1 is Normal-S1) and (Sen-2 is Rico-S2) then (output1 is CAT-S2) (1)
7. If (Sen-1 is Rico-S1) and (Sen-2 is Probre-S2) then (output1 is Reparar) (1)
8. If (Sen-1 is Rico-S1) and (Sen-2 is Normal-S2) then (output1 is S1) (1)
9. If (Sen-1 is Rico-S1) and (Sen-2 is Rico-S2) then (output1 is Reparar) (1)

Fig.9. Editor FIS en las reglas difusas para los sensores de oxígeno antes y después del catalizador.

1. If (TEMP is FRIO) and (RPM is RELENTI) then (output1 is NORMAL) (1)
2. If (TEMP is FRIO) and (RPM is NORMAL) then (output1 is NOHAYLECTURA) (1)
3. If (TEMP is TIBIO) and (RPM is RELENTI) then (output1 is NORMAL) (1)
4. If (TEMP is TIBIO) and (RPM is NORMAL) then (output1 is NORMAL) (1)
5. If (TEMP is NORMAL) and (RPM is RELENTI) then (output1 is NORMAL) (1)
6. If (TEMP is NORMAL) and (RPM is NORMAL) then (output1 is NORMAL) (1)
7. If (TEMP is CALIENTE) and (RPM is RELENTI) then (output1 is REPARAR) (1)
8. If (TEMP is CALIENTE) and (RPM is NORMAL) then (output1 is REPARAR) (1)

Fig.10. Editor FIS en las reglas difusas para los sensores de la temperatura refrigerante y las rpm

Para las reglas difusas fue necesario establecer las salidas difusas que son las fallas que se le pueden presentar al vehículo, el primer caso muestra las salidas difusas de los dos sensores de oxígeno (reparar, daño en el sensor 1, daño en el sensor 2) y daño en el catalizador con el sensor 2), y en el segundo caso se muestra las salidas difusa de la temperatura del líquido refrigerante con las rpm (buen funcionamiento, no hay lectura y reparar).

En la etapa de defuzzificación, se escogió la técnica del Centroide debido que una de las más utilizadas en el ámbito de la lógica difusa. En la Fig 11 y 12 se muestran las reglas If-



Then y se observa el método de defuzzificación empleado en el software Matlab.

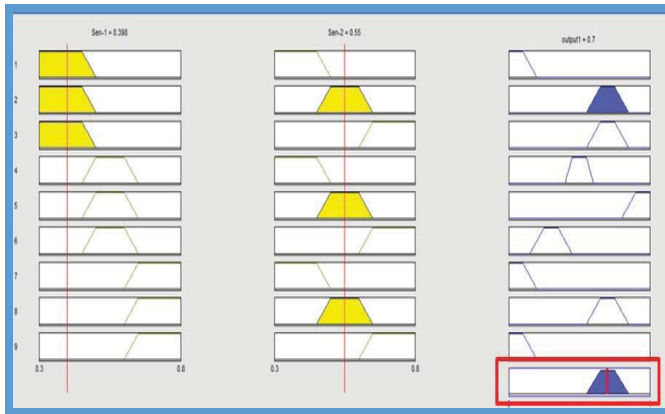


Fig.11. Reglas If-Then para los sensores de oxígeno antes y después del catalizador con el método de defuzzificación.

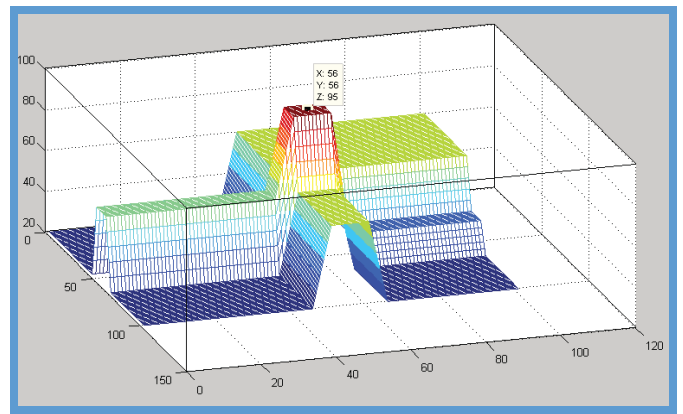


Fig.13. Distribución de salida de los sensores de oxígeno.

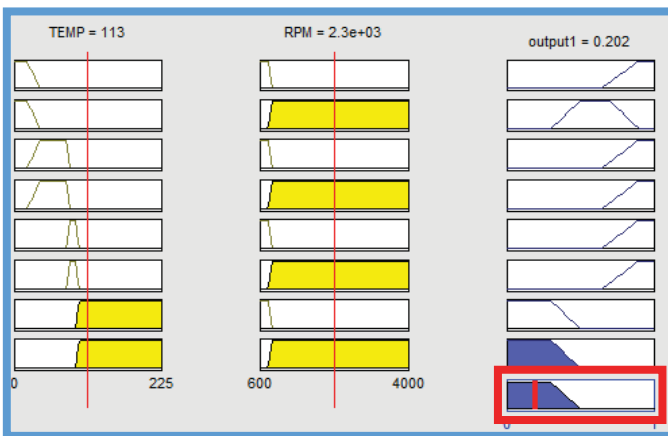


Fig.12. Reglas para los sensores de la temperatura refrigerante y las rpm con el método de defuzzificación.

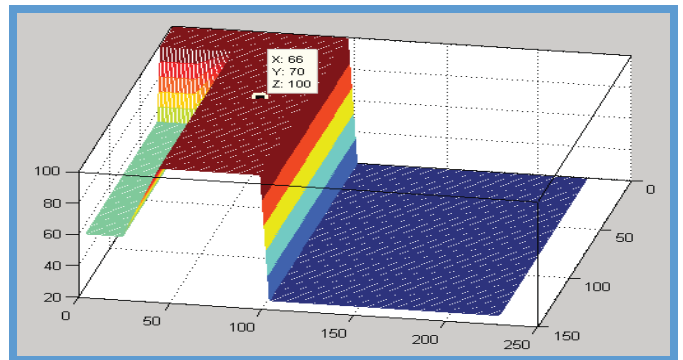


Fig.14. Distribución de salida de el sensor de temperatura y las rpm

#### IV. RESULTADOS

En la Fig.13 se muestra la distribución de salida para los sensores de oxígeno antes y después del catalizador donde se puede comprobar la efectividad de las reglas difusas el eje x muestra el valor del sensor 1, el eje y se muestra el sensor 2, y el eje z se observa el valor de la salida difusa o la posible falla que se puede presentar el vehículo. En la Fig. 14, se muestra el eje x el valor de la temperatura del líquido refrigerante, en eje y se muestra el valor de las rpm.

#### V. CONCLUSIONES

Los controles difusos de los sensores de oxígeno y el sensor de temperatura del líquido refrigerante con las rpm, presentan una salida difusa que permite inferir un diagnóstico de fallas de estos sensores, ya que esto le sirve al usuario a saber que avería tiene el vehículo, ya que algunas veces la computadora del carro no detecta la falla hasta el momento que afecte el buen funcionamiento del vehículo.

Los rangos de salida que se escogieron en los diferentes sistemas difusos fueron parte fundamental para obtener buenos resultados, debido a que la forma de organización de los conjuntos difusos de salida, se ven afectados por las reglas If- Then.

Como trabajo futuro se propone la implementación del algoritmo de lógica difusa presentado en este trabajo en un microcontrolador de la empresa Texas Instruments para evaluar el sistema en un contexto real y contemplar la adición de nuevas variables del vehículo al algoritmo, por ello se escogió la técnica del Centroide debido que es el más practico en la implementación en el hardware.

#### AGRADECIMIENTOS

Al Centro de Investigación en Ciencia Aplicada y Tecnología Avanzada – CICATA unidad Legaria, a la Escuela

Superior de Cómputo – ESCOM del Instituto Politécnico Nacional – IPN y al CONACYT por el apoyo, asesoría y acompañamiento brindado al desarrollo de este trabajo.

#### REFERENCIAS

- [1] I.C.Rodríguez, “los sensores automotrices en la practica” Vol.5, apendice 2, pp 2-6, 2011.
- [2] J. Lin, S.C. Chen, Y. T. Shih, y S. H. Chen, “A Study on Remote On-Line Diagnostic System for Vehicles by Integrating the Technology of OBD, GPS, and 3G”, World Academy of Science, Engineering and Technology International Journal of Social, Behavioral, Educational, Economic, Business and Industrial Engineering Vol:3, No:8, 2009.
- [3] M. Villeta, T. Lahera, S. Merino, J. G. Zato, J. E. Naranjo, y F. Jiménez, “Modelo para la Conducción Eficiente y Sostenible basado en Lógica Borrosa”, Revista Iberoamericana de Automática e Informática Industrial RIAI, Vol. 9, N°3, España, pp. 259-266, 2012.
- [4] V. Milanés, E. Onieva, J. Pérez, T. Pedro, y C. González, “Control de velocidad basado en lógica borrosa para entornos urbanos congestionados”, Revista Iberoamericana de Automática e Informática Industrial RIAI, Vol.6, N°4, España, pp. 61-68, 2009.
- [5] B. T. Matías, y D. I. Vicente, “Lógica borrosa,” Universidad Carlos III de Madrid, España, 2011
- [6] J. A. Cabrera, “Logica Clasica y logica clasica : facetas que las caracterizan”, Estoa, Universidad de Cuenca, ISBN: 1390-9274, Ecuador, pp. 91-101.
- [7] S. J. Hernández, “Aplicación de la lógica difusa para detección de defectos en rodamientos,” Departamento de Ingeniería Mecánica, Universidad Carlos III de Madrid, España, 2010.
- [8] M. J. A. Gallegos, M. A. C. Oviedo, O. C. Berdo, E. E. Ruiz, y M. A. L. Ortega, “Inteligencia artificial: modelos, técnicas y áreas de aplicación,” Editorial Paraninfo S.A, ISBN: 84-9732-183-9, España, pp. 84, 2003.
- [9] M. Cirstea, A. Dinu, M. McCormick, y J. G. Khor, “Neural and fuzzy logic control of drives and power systems. Newnes,” ISBN: 0-7506-55585, pp. 117-118, 2002.
- [10] C. Dualibe, M. Verleysen, y G.A Jaspers, “Design of Analog Fuzzy Logic Controllers in CMOS Technologies,” Kluwer Academic Publisher, The Netherlands, ISBN: 1-4020-7359-3, pp. 16-19, 2003.
- [11] P. Ponce, “Inteligencia artificial con aplicaciones a la ingeniería,” Alfaomega Grupo Editor, S.A, 1° Edición, ISBN: 978-607-7854-83-8, México, pp. 40-41, 2010.
- [12] B. Martín, y A. Sanz Molina, “Redes neuronales y sistemas difusos,” Alfaomega RA-MA Editorial, 2° Edición, ISBN: 97-0150-733-9, pp. 264-265, 2002.
- [13] F. A. M. Hernández, “Cooperación entre sistemas de inferencia, métodos de defuzzificación y aprendizaje de sistemas difusos lingüísticos,” Departamento de Ingeniería Electrónica, Sistemas Informáticos y Automática, Universidad de Huelva, España, 2008.
- [14] A. E. Cazar Ibarra, y S. X. Heredia Lema, “Diseño y desarrollo de un sistema experto para el mantenimiento preventivo y correctivo de las fallas más comunes en el sistema de inyección electrónica y sistema de escape de los motores a gasolina de cuatro tiempos de automóviles general motors, no de servicio público, que circulan en el distrito metropolitano de Quito, aplicando como operador de conocimiento para las conclusiones del análisis lógica Fuzzy,” Departamento de robótica e inteligencia artificial, Universidad Politécnica Salesiana, Quito, 2007.
- [15] S. N. Sivanandam, Sumathi, y S. N. Deepa, “Introduction to fuzzy logic using MATLAB,” Vol. 1. Springer, ISBN: 103-540-35780-7, Berlín, 2007, pp. 97-120.
- [16] M. A. Perez, y J. Verlon, “Sistemas de lógica difusa”. División de Electrónica y Computación CUCEI, Universidad de Guadalajara, México, 2008.
- [17] P. J. Escamilla, “Nuevo método para realizar el proceso de defuzzificación en controladores difusos,” Departamento de Electrónica, Laboratorio de Sistemas Neuro-Fuzzy, Universidad Nacional Autónoma de México, México, 2003.
- [18] M. Pérez, E. Cuevas, E., y D. Zaldivar, “Segmentación difusa. e-Gnosis,” Vol. 6, Universidad de Guadalajara, México, 2008.

# Propuesta de Metodología para la Construcción de Buscadores Web Semánticos de un Contexto Específico.

I.S.C. Jaime Martínez Vallejo  
Universidad Michoacana de San Nicolás de Hidalgo  
Morelia, México  
jaimemv99@hotmail.com

M.C. Juan Carlos Olivares Rojas  
Instituto Tecnológico de Morelia  
Morelia, México  
jcolivares@itmorelia.edu.mx

D.C. J. Guadalupe Ramos Díaz  
Instituto Tecnológico de La Piedad  
La Piedad, México  
jogar@hotmail.com

**Abstract**—El rápido crecimiento de la Web hace cada vez más difícil la recuperación de información a pesar de que los motores de búsqueda actuales implementan algoritmos muy sofisticados. Este trabajo propone una metodología que hace uso de las tecnologías de la Web Semántica para construir buscadores en un contexto determinado. La Web semántica está conformada por un conjunto de tecnologías, herramientas y estándares que conforman las bases de una infraestructura para soportar la visión de una Web asociada con significado. La metodología propuesta fue probada en contextos comerciales que proporcionan algún tipo de producto o servicio, como lo son restaurantes y hoteles.

La información inicial, que se utiliza como punto de partida, es extraída de un sitio Web con información temática o tipo directorio, a partir de la cual se puebla la ontología seleccionada, siendo ésta construida desde cero o bien se adapta de alguna ya existente, creando una base de datos semántica o data store; por último, se desarrolló una aplicación de búsqueda, que mediante los criterios indicados, busca y muestra la información encontrada. Los resultados demuestran que un buscador que utiliza las tecnologías de Web Semántica puede obtener buenos resultados comparándolo con los motores de búsqueda tradicionales, en los casos de contextos muy específicos.

**Keywords**—Web Semántica, Buscadores, Ontología Hoteles, Ontología Restaurantes

## Introducción

Desde sus orígenes, la Web ha venido revolucionando la manera en que la gente se comunica y lleva a cabo una gran cantidad de actividades; se ha convertido en el corazón de una revolución que está transformando el mundo desarrollado hacia lo que se conoce como la Sociedad del Conocimiento [1]. En 25 años la Web ha transformado al mundo y se ha vuelto indispensable, impactando la actividad humana en

prácticamente todos los campos. Este crecimiento de la Web, plantea el reto de la encontrar la información requerida por el usuario, ya que este continuo y rápido crecimiento en el volumen de la información dificulta su búsqueda, organización, acceso y mantenimiento [2].

A pesar de que la tecnología de los motores de búsqueda ha ido mejorando, aún persisten dificultades para encontrar información, tal parece que la cantidad de contenido Web supera el avance tecnológico. [1]

## Web Semántica

La Web Semántica se puede definir como una colección de estándares y tecnologías que permiten a las máquinas (por medio de programas), interpretar (entender) el significado (semántica) de la información en la Web. Es una extensión de la Web actual, en la cual la información se representa con un significado bien definido, permitiendo que las computadoras y la gente trabajen de manera cooperativa.

En [3] se menciona que algunos de los problemas de la Web actual en la recuperación de información mediante los motores de búsqueda tradicionales que se pueden resolver con la Web Semántica son la ambigüedad en los resultados, que se genera debido al proceso de indización en texto plano utilizado por los buscadores y el poder realizar búsquedas complejas, así como descargar al usuario la actividad de revisar manualmente la gran cantidad de resultados devueltos.

La visión de la Web Semántica del consorcio W3C está comprendida en cuatro componentes principales:

- 1.- Expresión del significado.
- 2.- Representación del conocimiento.
- 3.- Ontologías.
- 4.- Agentes.

La expresión del significado es fundamental para la construcción de la Web Semántica. La Web actual carece de mecanismos para expresar el significado y por lo tanto es estática. La representación del conocimiento proporciona el mecanismo que permite expresar el contenido en un formato estructurado para que sean aplicados métodos de inferencia para obtener conclusiones útiles. Para lograr que la representación del conocimiento sea con significado y a la vez práctica, el significado detrás de los datos debe ser compartido. Esto se puede lograr utilizando ontologías, las cuales se refieren a un vocabulario compartido acerca de algún dominio o concepto. La premisa es que si ese vocabulario es compartido con respecto a un concepto, entonces el significado asociado al concepto es claro entre los que comparten el vocabulario. Una vez que la ontología se ha acordado por una comunidad y si la ontología puede ser representada en un formato entendible entonces se pueden utilizar agentes de software para razonar con el conocimiento representado y capturado mediante dicha ontología.

**Objetivo.**

Este trabajo pretende desarrollar una propuesta metodológica para construir un buscador Web semántico de un contexto dado, a partir de la definición de criterios de búsqueda y que obtenga resultados precisos.

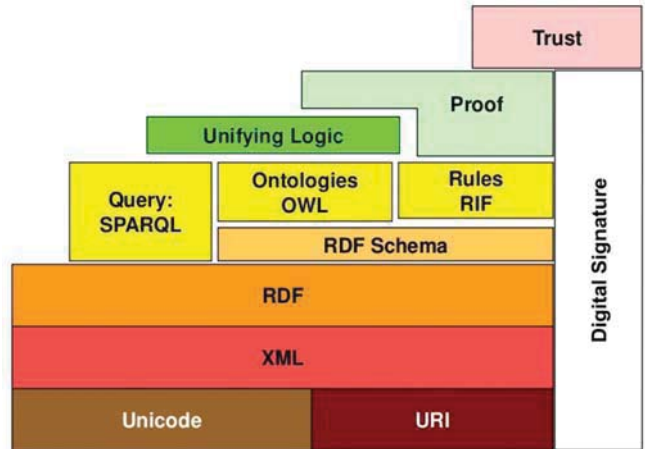
Además, se pretende demostrar que la implementación de las tecnologías de Web Semántica en aplicaciones de búsqueda de información para un contexto específico, ofrecen ventajas frente a los motores de búsqueda tradicionales, ya que aun cuando los principales motores de búsqueda actuales son bastante útiles, aun se presentan casos de ambigüedad en las búsquedas y no existen condiciones favorables para que se puedan realizar procesos de búsqueda, extracción e integración de información de forma sistematizada.

**I. MARCO TEÓRICO**

**Componentes de la Web Semántica**

La Web semántica está conformada por un conjunto de tecnologías, herramientas y estándares que conforman las bases de una infraestructura para soportar la visión de una Web asociada con significado. La arquitectura de la Web semántica está compuesta de una serie de estándares organizada en una cierta estructura que es una expresión de sus interrelaciones. Esta arquitectura frecuentemente se representa utilizando un diagrama propuesto por vez primera por Tim Berners-Lee. La figura 1 muestra las partes de la arquitectura de la Web semántica. Comienza con la base de URIs (Uniform Resource Identifier- Identificador Uniforme de Recursos) y Unicode. Encima de esto encontramos la capa de interoperatividad sintáctica en la forma de XML (eXtensible Markup Language- Lenguaje Extensible de Marcado), que a su vez soporta a RDF (Resource Description Framework-Estructura de descripción de Recursos) y RDF-S (RDF Schema- Esquema RDF). Los lenguajes de Ontología Web se construyen sobre RDF(S). Las últimas tres capas son la de lógica, prueba y confianza, que hasta la fecha no han sido explotadas de manera significativa.

Figura 1. Diagrama de Capas de Web Semántica



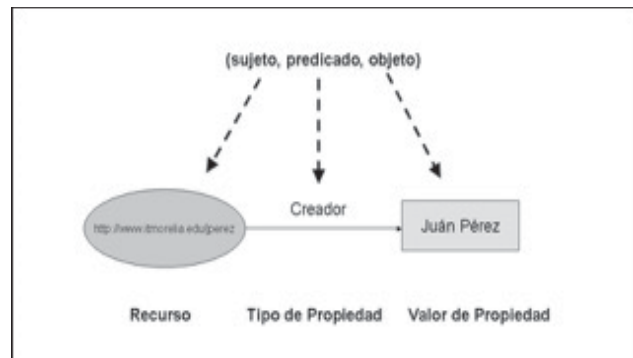
RDF es un sistema en el que se describe un recurso por medio de una serie de propiedades relacionadas. Cada descripción de propiedad se compone de una sentencia, que a su vez se compone de un tipo de propiedad, un valor y el identificador del recurso sobre el que se aplica la propiedad. Utilizando este modelo, cualquier recurso que posea un identificador uniforme de recursos (URI) se puede describir por medio de RDF. [4]

La siguiente sentencia legible para una persona:

**Juan Pérez es el creador de la Página Personal de Juan Pérez**

se puede representar gráficamente mediante gráficos etiquetados, tal como se muestra en la figura 2. La página personal de Juan Pérez es un recurso. Este recurso tiene la URI <http://www.itmorelia.edu/jperez>, y tiene la propiedad creador, con el valor "Juan Pérez".

Figura 2. Representación sentencia RDF.



**OWL**

OWL es el acrónimo de Lenguaje Web de Ontologías (Web Ontology Language), es un lenguaje RDF desarrollado por el W3C para definir clases y propiedades, y también para proporcionar mayor capacidad de razonamiento e inferencia sobre las relaciones. OWL se construyó como una extensión de RDFS, y es el estándar actual del W3C para definir esquemas de

Web Semántica, además de herramientas y soporte API para OWL cuyo uso se ha expandido rápidamente [5].

Tecnologías de gestión y consulta

#### SPARQL

Es un lenguaje de consulta RDF y protocolo de acceso de datos para la Web Semántica. Su nombre es un acrónimo recursivo de SPARQL Protocol and RDF Query Language. Fue estandarizado por el Grupo de Trabajo SPARQL del W3C en enero del 2008.

#### RDF Data Store

RDF Data store y RDF Database store son sinónimos y se utilizan de manera intercambiable. Es una base de datos de propósito especial construida para almacenar y recuperar sentencias RDF.

#### JENA

JENA es un framework JAVA que proporciona un entorno de programación para los estándares de Web Semántica RDF, RDFS, OWL y SPARQL e incluye un motor de inferencia basado en reglas.

## II. TRABAJOS RELACIONADOS

Existen una serie de trabajos relacionados como: Generador Semiautomático de Perfiles de usuario mediante OWL [6], el cual implementa una propuesta de Modelado del perfil de usuario mediante el uso de ontologías, para que pueda ser utilizado en sistemas de recomendación contextuales. Uno de los objetivos es minimizar la intrusión en el proceso de extracción de información personal del usuario. Se analizaron las ontologías existentes Aktors, FOAF y SOUPA, para seleccionar atributos de algunas de ellas. Como producto se obtiene una ontología con gran capacidad de explotación por parte de servicios o sitios Web orientados a los contenidos.

Desarrollo de una Web API para el tratamiento Automático de Reglas Contextuales Aplicadas a Servicios Basados en Localización [7], el cual presenta el desarrollo de una API, que implementa tecnologías de Web Semántica para recomendar servicios en el dominio de hospedajes. Utiliza servicios basados en localización para ubicar hoteles cuyas características modeladas en una ontología y contrastarlas con las propiedades de perfiles de usuario y ofrecer los servicios de hospedaje coincidentes, tanto con el perfil de usuario como con variables de contexto.

Búsquedas contextuales de Servicios Basados en Localización en un entorno de Web social [8], que presenta la propuesta de un sistema de recomendación por filtrado colaborativo utilizando ontologías y reglas de Web Semántica, que permiten representar el modelo de usuario y su contexto. Hace uso de anotaciones sociales e información contextual, además de la teoría de la mercadotecnia.

Poblado semiautomático de Ontologías Organizacionales a partir de Análisis de contenido Web [9], propone la extracción de información de un sitio Web utilizando un conjunto de reglas y especificaciones para que a partir de la extracción se realice un

poblado ontológico de un dominio específico. El proceso propuesto consta de las etapas de Recuperación de Información, Extracción de Información y Poblado Ontológico.

Effects of relevant contextual features in the performance of a restaurant recommender system [10], propone la utilización de información contextual para mejorar la satisfacción del usuario en sistemas de recomendación. El trabajo se centra en la selección de atributos de contexto para mejorar la habilidad predictiva de un sistema de recomendación. El estudio se centra en Surfeous, un sistema prototipo para recomendación basado en filtros colaborativos y modelos semánticos. Se identificaron variables relevantes aplicando un enfoque de selección de criterios. La investigación tiene impacto en los aspectos de identificación de atributos contextuales relevantes que los usuarios toman en cuenta al seleccionar un restaurant, reducción de la dimensionalidad y proporciona nuevas percepciones sobre los efectos de las variables contextuales en la predicción del desempeño de sistemas de recomendación.

Especificación de una Ontología para teleeducación en la Web Semántica [11], se especifica una ontología completa en el dominio de la teleeducación, desarrollada en el lenguaje OWL. La ontología modela conceptos relacionados a la interacción de los distintos agentes que intervienen en una experiencia de aprendizaje a través de la Web. La arquitectura propuesta permite que agentes inteligentes puedan buscar información semántica necesaria para encontrar fuentes de objetos educativos, así como descargarlos de los servidores correspondientes.

## III. METODOLOGÍA PARA LA CONSTRUCCIÓN DE BUSCADORES WEB SEMÁNTICOS DE UN CONTEXTO ESPECÍFICO

En este apartado se describe la metodología propuesta para la generación de un sistema de búsqueda Web Semántica en un contexto o dominio específico a partir de la información de uno o varios sitios web temáticos sobre dicho contexto o dominio.

El contexto o dominio puede ser sobre cualquier temática, para probar la metodología se aplicaron ejemplos en entidades comerciales que proporcionan algún tipo de producto o servicio, como los restaurantes y los hoteles.

La información inicial, que se utiliza como punto de partida, es extraída de un sitio Web con información temática o tipo directorio, a partir de la cual se puebla o llena la ontología seleccionada, siendo ésta construida de cero o bien adaptar alguna ya existente, creando una base de datos semántica o data store; por último se desarrolla una aplicación de búsqueda, que mediante los criterios indicados, busca y muestra la información encontrada. La figura 3 muestra la metodología en general.

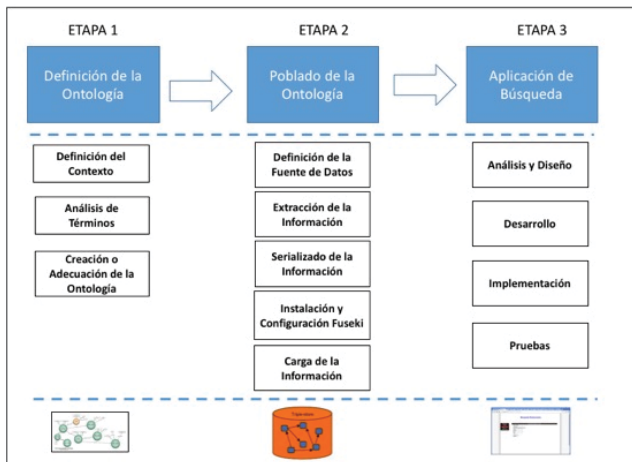


Figura 3. Etapas de la metodología propuesta.

Etapa 1: Definición de la Ontología del Contexto. En esta etapa se analiza el contexto específico sobre el que se desarrollará el buscador, para llegar a definir la ontología que será utilizada en la representación de la información semántica, definiendo de manera precisa las clases, atributos y relaciones. Se utiliza la metodología sugerida por [12], para el desarrollo de una ontología, la cual es una de las más completas según se menciona en [13].

Para el caso del contexto de restaurantes se adaptó la ontología Goodrelations [14] y para el contexto de hoteles se adaptó la ontología Accomodation [15] ya que al aplicar los pasos sugeridos por [6] fue la opción más adecuada. En la figura 4 se muestra el diagrama de clases de la ontología restaurantes para el contexto correspondiente.

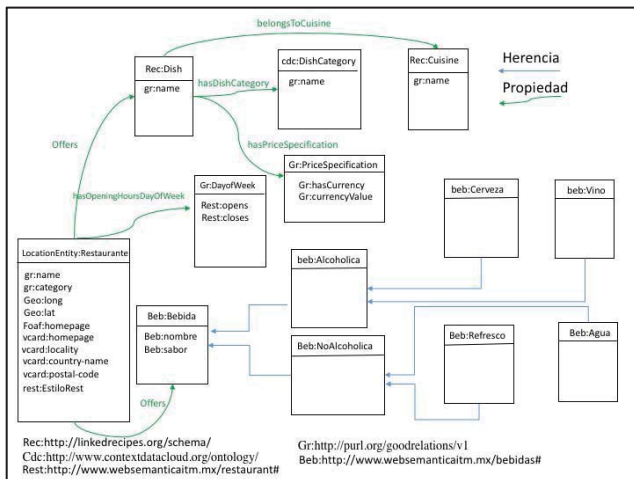


Figura 4. Ontología Restaurant

Etapa 2: Poblado de la Ontología. Aquí se realiza la extracción de la información a partir del sitio o sitios seleccionados, dicha información se lleva a una estructura de tabla y a partir de ésta se serializa en formato RDF/XML, para crear las tripletas correspondientes a la ontología. La ontología poblada queda almacenada en la base de datos RDF Fuseki [16]. Para completar esta etapa se realizan los procesos de definición

de el origen de datos, la extracción de la información, el serializado, la instalación y configuración del Triple Store Server y la carga de información.

Etapa 3: Desarrollar aplicación Buscador. En esta etapa se definen los criterios de búsqueda, así como el desarrollo de aplicación de búsqueda, la cual a partir de los criterios proporcionados realizara una consulta a la base de datos RDF, utilizando el lenguaje de consultas SPARQL y haciendo uso de la librería JENA, para después desplegar los resultados obtenidos. Esta aplicación se desarrolló con la tecnología Java Servlet/JSP. En la figura 5 se muestra una imagen de la pantalla de búsqueda de restaurantes. Para completar esta etapa se deben realizar los procesos de análisis y diseño, desarrollo, implementación y pruebas de la aplicación de búsqueda.



Figura 5. Pantalla de Búsqueda restaurantes.

Como se puede observar en la figura 3, al finalizar cada una de las etapas se obtiene un producto, en la primera etapa se define la Ontología, en la segunda etapa se obtiene la base de datos semántica, y en la tercera etapa el producto es la aplicación de búsqueda.

#### Validación de la Metodología

Aunque no se define un procedimiento para validar la metodología, la adecuada aplicación de cada una de las etapas garantiza el éxito en la implantación del buscador mediante Web Semántica.

## IV. PRUEBAS Y RESULTADOS

### Pruebas de uso de la metodología.

La metodología propuesta se aplicó en dos contextos específicos, el primero de ellos fue el de Restaurantes y el segundo el de Hoteles, en ambos casos se logró obtener un Buscador Web Semántico para cada dominio o contexto.

A continuación se presenta para cada uno de los casos los productos obtenidos en cada una de las etapas.

### Dominio Restaurantes.

### Ontología de Restaurantes.

Se definió a partir de la Ontología GoodRelations [14], ya que es una de las más completas, la adecuación específica consistió en agregar clases para la presentación de bebidas.

La clase Restaurante es una subclase de LocationEntity, y mediante la propiedad Offers, ofrece platillos (clase Dish) y bebidas (clase Bebidas).

### Triple Store para Restaurantes.

La base de datos semántica generada a partir del Archivo RDF, contiene 782 tripletes de las instancias de 10 restaurantes y la definición de las clases y propiedades contiene 56 tripletes, las cuales se almacenaron en el triple store Fuseki [16].

### Buscador de Restaurantes.

Para el buscador se definieron las palabras clave de búsqueda Tipo de Cocina, Tipo de Alimento o Nombre de Platillo, Bebida y horario, tal como se muestra en la figura 5. Por decisión de diseño se definieron estos campos como criterios de búsqueda, aunque se pueden utilizar cualquiera de las clases o propiedades definidas en la ontología. La búsqueda se realiza tomando los cuatro criterios y se presentan como resultados todas las combinaciones de platillo/bebida que coincidan con los valores indicados, además de incluir el nombre del restaurante cuyo horario de atención cubra el día y hora que se indica en la consulta. Se mostrarán solo los resultados en los que coincida la totalidad de los criterios, es decir, si existe un platillo y bebida en los parámetros de búsqueda que coincida con algún restaurante, pero el tipo de cocina de dicho restaurante no coincide, no aparecerá dicho resultado.

### Dominio Hoteles.

### Ontología de Hoteles.

Se utilizó la ontología Accomodation [14], la cual es una ontología que permite representar varios tipos de alojamiento, como son hoteles, cabañas, sitios para acampar, etc. Para alcances de uso de la metodología se utilizó solo la parte de hoteles. Las clases principales que se utilizaron son la clase hotel y la clase habitación (Room) la cual se relaciona mediante la propiedad *partOf* (parte de) con la clase hotel.

### Triple Store para Hoteles.

La base de datos semántica generó 21 tripletes en la definición de la clase y 1590 para la información de instancias de diez hoteles. El sitio Web que se tomó de referencia fue zonaturistica.com filtrando datos de hoteles en Morelia.

### Buscador de Hoteles.

Se utilizó como campos para criterios de búsqueda el cupo máximo requerido para la habitación, servicios o facilidades de la habitación, el precio máximo por noche y la categoría del hotel.

Para los resultados de búsqueda se incluyó el nombre del hotel con sus datos generales, el precio de la habitación por noche, el nombre o descripción de la habitación, el servicio

buscado y una foto de la habitación, en el caso de este ejemplo, el parámetro categoría de hotel es opcional y el resto obligatorios, es decir, aparecerán resultados que coincidan con los primeros tres parámetros pero no necesariamente con el cuarto, que es la categoría.

Lo anterior es una decisión de diseño con el objetivo de demostrar que los criterios de búsqueda pueden ser variados. En la figura 6 se muestra una pantalla con los resultados de una búsqueda de hoteles.

Hotel	Precio Día	Habitación	Servicios	Foto
POSADA VISTA BELLA 4 Entradas <a href="http://morelia-hotels.com/maestros-hoteles/posada-vista-bella">http://morelia-hotels.com/maestros-hoteles/posada-vista-bella</a>	576	Estandar	Desayuno Buffet, Mesa de trabajo	
HOTEL HOWARD JOHNSON CALLE REAL 4 Entradas <a href="http://www.hjmorelia.com">www.hjmorelia.com</a>	740	Ejecutiva II	Mesa de Trabajo	

Figura 6. Resultados de Búsqueda de Hoteles

### Pruebas del Buscador Web Semántico.

### Pruebas de comparación del Buscador Semántico.

Se llevaron a cabo pruebas de comparación entre el buscador propuesto y los motores de búsqueda Google, Yahoo e Ixquick, realizando algunas consultas en cada uno de ellos y comparando la relevancia de los resultados obtenidos.

### Consideraciones.

Para llevar a cabo pruebas comparativas entre el buscador de la metodología propuesta y los buscadores tradicionales se debe considerar:

- Se definieron preguntas para realizar búsquedas y llevar a cabo pruebas comparativas entre los buscadores de google, yahoo, ixquick y la propuesta presentada.
- Dado que los buscadores tradicionales realizan sus búsquedas a partir de una cadena de búsqueda y el buscador propuesto lo hace a partir de parámetros, cada una de las preguntas se adecuó bajo el siguiente formato:

Para el contexto de Restaurantes:

Pregunta en lenguaje Natural: Restaurantes en Morelia de Cocina Mexicana que ofrezcan platillos con pollo, que tengan cerveza y que estén abiertos a las 12:00 pm.

Texto de consulta en motores de búsqueda:  
Restaurantes Morelia cocina mexicana pollo cerveza abiertos a las 12:00

Parámetros en Buscador Semántico: Mexicana, Pollo, Cerveza,12:00

Para el caso del Buscador semántico se omitieron los términos Restaurante y Morelia, ya que el contexto es sobre restaurantes y la aplicación de prueba se cargó con información de restaurantes de Morelia únicamente, para efectos de pruebas.

• Para cada búsqueda realizada se analizaron los primeros 5 resultados devueltos por cada motor de búsqueda, asignando para efectos prácticos un valor de relevancia a cada resultado que va desde 0 hasta 1, aplicando la siguiente métrica, basada en [17] para el cálculo de dicha relevancia:

$$\text{Relevancia} = \text{Valoración Contexto} + \text{Valoración Precisión}$$

Valoración para contexto.

Valoración para contexto	
Valoración	Descripción
0.0	Fuera de contexto. El resultado no tiene que ver con información del dominio específico. (Restaurantes)
0.5	Dentro del Contexto. El resultado tiene que ver con información del dominio específico en la ciudad buscada (Restaurantes en Morelia).

Valoración Precisión (Acumulable)	
Valoración	Descripción
0.125	El resultado incluye el parámetro de búsqueda Tipo de Cocina.
0.125	El resultado incluye el parámetro de búsqueda Tipo de Alimento/Nombre de platillo.
0.125	El resultado incluye el parámetro de búsqueda Bebida.
0.125	El resultado incluye el parámetro de búsqueda hora de servicio.

Una vez que se analiza cada resultado se le asigna la valoración correspondiente tanto en contexto como en precisión y para obtener la valoración total de relevancia se suman los valores obtenidos, lo que devolverá un valor entre 0 y 1 para cada resultado.

• Se omitieron los resultados duplicados, ya que en varios casos con los motores de búsqueda tradicionales apareció más de una vez la misma página.

• Para el buscador propuesto los resultados son por cada platillo y bebida que ofrezca un restaurante, por lo que si existen dos platillos a base de pollo en el mismo restaurant, aparecerán dos resultados. Para efectos de esta comparativa se considera este caso como un único resultado.

#### Resultados

Se definieron las preguntas y se tuvieron los siguientes resultados:

Pregunta 1: Restaurantes en Morelia de Cocina Mexicana que ofrezcan platillos con pollo, que tengan cerveza y que estén abiertos a las 12:00 pm.

Resultados Contexto Restaurantes		
Valoración	Total Resultados	Relevancia (Promedio 5 Primeros)
Google	2,150	72%
Yahoo	31,600	48%
Ixquick	No Especificado	65%
Propuesta Buscador Semántico	2	100%

Pregunta 2: Hoteles en Morelia con habitación con cupo máximo de dos personas, que ofrezcan servicio de café en la habitación, con un costo por noche de \$800.00 máximo y de categoría 4 estrellas.

Pregunta en Motores de búsqueda: “hoteles Morelia 4 estrellas habitación 2 personas servicio de café 800 por noche”

Resultados Contexto Hoteles		
Valoración	Total Resultados	Relevancia (Promedio 5 Primeros)
Google	7,100	49%
Yahoo	86,600	48%
Ixquick	No Especificado	65%
Propuesta Buscador Semántico	3	96%

#### Conclusiones de los resultados.

Los resultados nos muestran que el buscador propuesto es efectivo en cuanto a la relevancia definida en las pruebas comparativas, aunque la cantidad de información recuperada es pequeña en relación a los motores de búsqueda tradicionales. Es importante mencionar que las decisiones del diseño del buscador en cada uno de los ejemplos impactan directamente en los resultados del mismo, ya que si se establecen la totalidad de los criterios de búsqueda como requisitos incluyentes en la búsqueda, se garantiza que la relevancia de los resultados obtenidos sea del 100%, tal como se observa en el ejemplo de los restaurantes donde los cuatro parámetros se deben cumplir para que aparezcan en el listado de resultados, a diferencia del caso de los hoteles donde el parámetro de categoría es opcional que se cumpla, lo que arroja valores de relevancia menores al 100% para los casos en que ese criterio no se cumpla.

Se observó que los motores de búsqueda tradicionales logran un número muy alto de resultados, no todos relevantes, mediante la búsqueda por ocurrencia de palabras. Esto ocasiona que el usuario deba invertir tiempo y recursos en revisar los resultados y localizar información precisa y útil.

## V. CONCLUSIONES Y TRABAJOS FUTUROS

### Conclusiones

La metodología propuesta en el presente trabajo cumple con su principal objetivo que consiste en construir un Buscador Web Semántico para un contexto específico a partir de la información de un sitio Web tradicional de tipo directorio.

La metodología comprende las etapas de Definición de la Ontología, Poblado de la Ontología y Desarrollo de la



Aplicación, dichas etapas conforman una estructura modular, en la que se puede modificar o mejorar alguna de ellas sin afectar de manera sustancial la metodología.

La aplicación de la metodología en dos dominios o contextos diferentes asegura la factibilidad de la misma para cualquier otro dominio, siendo la primera etapa la que puede requerir un análisis muy particular dependiendo del contexto que se vaya a utilizar.

La segunda etapa es susceptible de mejorarse sobre todo en la parte de la extracción de la información del sitio Web que se utilice como fuente de información y en la tercera etapa, que es el desarrollo de la aplicación se presentaron ejemplos sencillos para efecto de prueba de la metodología.

En el apartado siguiente se sugieren mejoras en algunos puntos de las etapas que pueden perfeccionar y ampliar la efectividad de la metodología propuesta.

#### Trabajos futuros.

Los trabajos futuros factibles para complementar la metodología son:

Implementación de mecanismos que automaticen la extracción de información del sitio Web origen, mediante el uso de herramientas de análisis y manipulación de documentos HTML, para que identifique la estructura del sitio y se implementen rutinas para la extracción automática.

- Aplicación de técnicas para el Procesamiento de lenguaje natural que permitan términos de búsqueda amigables para los usuarios y no esté limitado a utilizar un número de parámetros de búsqueda fijos.

- Mecanismos para resolver el uso de sinónimos en las búsquedas, de manera que funcionen aunque en el texto de búsqueda se utilicen términos equivalentes a los definidos en la ontología.

- Mecanismos para gestionar búsquedas en idiomas diferentes al idioma en que esté definida la Ontología

#### REFERENCIAS

- [1] Antoniou, G., & van Harmelen, F. (2008). *A Semantic Web Primer* (2da. ed.). Cambridge: The MIT Press.
- [2] Davies, J., Fensel, D., & van Harmelen, F. (2003). *Towards the Semantic Web*. London: Jhon Willey & Sons, LTD.
- [3] Domingue, J., Fensel, D., & Hendler, J. A. (2011). *Handbook of Semantic Web Technologies*. (J. Domingue, D. Fensel, & J. A. Hendler, Edits.) Berlin: Springer-Verlag.
- [4] Cardoso, J. (Sept/Oct de 2007). *The SemanticWeb Vision: Where Are We?* (IEEE, Ed.) *IEEE Intelligent Systems*, 22-26.
- [5] Segaran, T., Evans C. & Taylor, J. *Programming The Semantic Web*. Sebastopol: O'Reilly Media, Inc.
- [6] Rojas Roldán, Christian E.. (2009). *Generador automático de perfiles de usuario mediante Owl*. Cuernavaca, Mexico. Cenidet.
- [7] Villatoro H, César Francisco. (Febrero 2011). *Desarrollo de una Web API para el tratamiento Automático de Reglas Contextuales Aplicadas a Servicios Basados en Localización*. Cuernavaca, México. Cenidet.
- [8] Ponce Medellín Israel R. (2011). *Búsquedas contextuales de Servicios Basados en Localización en un entorno de Web social*. Cuernavaca, Mexico. Cenidet.

- [9] Munguía Aguilar Everardo. (2012). *Poblado semiautomático de Ontologías Organizacionales a partir de Análisis de contenido Web*. Cuernavaca, Mexico. Cenidet.
- [10] Vargas-Govea Blanca, Gonzalez-Serna Gabriel, Ponce-Medellin Rafael (2011). *Effects of relevant contextual features in the performance of a restaurant recommender system*. *Proceedings of 3rd Workshop on Context-Aware Recommender Systems* (2011).
- [11] Romero-Loop Roberto (2007). *Especificación de una Ontología para teleeducación en la Web Semántica*. Valencia, España. Universidad Politécnica de Valencia.
- [12] Noy, N. F., & McGuinness, D. L. (s.f.). *protege*. Recuperado el 12 de 11 de 2015, de *protege*: [protege.stanford.edu/publication/ontology\\_development/ontology101.pdf](http://protege.stanford.edu/publication/ontology_development/ontology101.pdf)
- [13] Guzmán Luna, J. A., Lopez Bonilla, M., & Durley Torres, I. (2012). *Metodologías y métodos para la construcción de ontologías*. (U. T. Pereira, Ed.) *Scientia et Technica*, XVII (50).
- [14] Hepp, Martin. *GoodRelations Ontology*, recuperado el 12 de Octubre de 2015. <http://www.heppnetz.de/projects/goodrelations/>.
- [15] Hepp, Martin. *Accommodation Ontology Metadata*, recuperado el 12 de Octubre de 2015. <http://purl.org/acco/ns>. Innsbruck University.
- [16] Apache Fuseki. Recuperado el 12 de Octubre de 2015. [https://jena.apache.org/documentation/serving\\_data/](https://jena.apache.org/documentation/serving_data/). Apache Coprp.
- [17] Javed Ahmad Khan, Deepak Sangroha, Masroor Ahmad, Md.Tanzillur Rahman. *A Performance Evaluation of Semantic based Search Engines and Keyword based Search Engines* (2014). *International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom)*.
- [18] Hebel, J., Fisher, M., Blace, R., & Perez-Lopez, A. (2009). *Semantic Web Programming*. Indianapolis: Willey Publishing, Inc.
- [19] Murguía Aguilar, E. (2012). *Poblado Semiautomático de Ontologías Organizacionales a partir de Análisis de Contenido Web*. (Cenidet, Ed.) Cuernavaca, Morelos, Mexico.
- [20] *OWL Web Ontology Language Use Cases and Requirements*. (Febrero de 2004). Obtenido de W3C: <http://www.w3.org/TR/webont-req/#onto-def>
- [21] Passin, T. (2004). *Explore's Guide to the Semantic Web*. Greenwich: Manning Publications.
- [22] Ramos Diaz, J. (s.f.). *Informe de año sabático*. AS-1-101/2014.
- [23] Sikos, L. F. (2015). *Mastering Structured Data on the SEMantic Web*. Apress.
- [24] Staab, S., & Studer, R. (2009). *Handbook of Ontologies* (2da. ed.). New York: Springer.
- [25] Yu, L. (2011). *A Developer's Guide to the Semantic Web*. Heidelberg: Springer.

# Aplicación para el registro de resultados de estudios de Imagenología dentro del Expediente Clínico Electrónico mediante HL7

I.S.C. Nancy Aracely Cruz Ramos  
*Instituto Tecnológico de Orizaba*  
 Veracruz, México  
 ncruzamos@acm.org

M.C.E. Beatriz Alejandra Olivares Zepahua  
*Instituto Tecnológico de Orizaba*  
 Veracruz, México  
 bolivares@ito-depi.edu.mx

**Resumen** - En este trabajo se presenta una aplicación Web para la integración de los resultados de estudios de Imagenología dentro del Expediente Clínico Electrónico (ECE) de un sanatorio. El modelado de la arquitectura consideró la Norma Oficial Mexicana NOM-024-SSA3-2012 y el estándar de mensajería HL7; también incluye una ontología de conceptos médicos, la cual es necesaria para la construcción de mensajes HL7 de manera automatizada y la comunicación entre el sistema de ECE y los equipos médicos que soporten la versión 2.x de HL7.

**Palabras clave:** Expediente Clínico Electrónico, Imagenología, HL7, ontología, equipo médico.

## I. INTRODUCCION

En la actualidad diversas empresas a nivel mundial se dedican al desarrollo de Sistemas de información en el área de Salud, específicamente Sistemas de Expediente Clínico Electrónico (ECE). Sin embargo, en México estas aplicaciones aún se encuentran en fase de implementación y, en muchos casos, no contemplan la integración de los resultados de estudios de Auxiliares de Diagnóstico como Imagenología en el expediente del paciente.

La información de dichos estudios se almacena en diferentes formatos y plataformas informáticas y se encuentra directamente en los equipos médicos, es decir, no se vincula a un ECE lo que dificulta las búsquedas y provoca un retraso significativo en el diagnóstico médico [1].

A diferencia de las aplicaciones existentes con características similares a la aplicación que se desarrolló en este trabajo, ésta funciona con cualquier tipo de equipo médico del área de Imagenología y permite la construcción automática de mensajes HL7 (necesarios para comunicarse con los equipos médicos), esto se logra a partir de consultas a una ontología de conceptos médicos que fue creada en este proyecto.

## II. MARCO TEÓRICO

### 2.1 Estudios de Auxiliares de Diagnóstico

Estos estudios son exámenes que sirven a un médico para apoyar o descartar un diagnóstico y reportan el comportamiento de parámetros específicos del organismo o el estado del mismo mediante su apreciación sensorial o visual.

Entre los estudios de Auxiliares de Diagnóstico más importantes se encuentran: estudios de Laboratorio, estudios de Imagenología, electrocardiograma, encefalograma, endoscopia y termografía.

Como los instrumentos científicos cada vez tienen mayor sofisticación y precisión, surgieron las áreas Auxiliares de Diagnóstico que aportan datos e imágenes que se relacionan con estados de salud o enfermedades.

La integración de los estudios Auxiliares de Diagnóstico en la práctica médica es importante ya que interviene en el seguimiento del tratamiento y en área medicina preventiva. Estos estudios proporcionan información para complementar la que el médico a través de sus conocimientos y experiencia establece en la consulta [2].

### 2.2 Mensaje HL7

HL7 es una de las organizaciones más importantes en estándares de mensajería en informática médica. Los mensajes HL7 permiten el intercambio de datos entre sistemas informáticos de salud. Para transferir la información, la sintaxis de los mensajes utiliza cadenas ASCII con delimitadores [3]. En la Fig. 1 se muestra un ejemplo de mensaje HL7 V2.x.

```
MSH|^~\&|GI|1000||1000|20151208||ORM^001|1|T|2.5
PID|1||P0001||Huerta Pliego^Antonio||19570722|M
PV1|0001|E|URG^URG-05^URG-05|||||ENFERMERIA
ORC|NW|1|||||20151208
OBR|1|1||2080006723^TOMOGRAFIA DE CRANEO||20151208
```

Fig. 1 Ejemplo Mensaje HL7 con la sintaxis de la versión 2.x

Los mensajes HL7 se dividen en segmentos y cada segmento en campos. Cada campo HL7 esta integrado por componentes y a cada componente pertenece un tipo de dato específico. Los segmentos más relevantes de un mensaje HL7 V2.x son los siguientes:

MSH → Cabecera del mensaje.  
 PID → Identificación del paciente.  
 PV1 → Información acerca del episodio.  
 ORC → Datos generales de la solicitud.  
 OBR → Detalle de solicitud.

### 2.3 Mirth Connect

Mirth Connect es un motor de integración y software basado en Java, que hace uso de la biblioteca HAPI y se diseñó específicamente para trabajar con datos en formato HL7.

Mirth Connect permite la emisión bidireccional de mensajes HL7 entre sistemas y aplicaciones sobre múltiples capas de transporte y su arquitectura se basa en canales de información de forma que permite el filtrado de mensajes, transformación y enrutamiento de los mismos con base en reglas que define el usuario. Los canales son terminales (de entrada y de salida), filtros y transformadores. Las terminales se utilizan para la configuración de las conexiones y los protocolos. Las terminales de entrada se utilizan para designar la forma en que llegan los mensajes de entrada, por ejemplo TCP/IP o un servicio Web. Las terminales de salida se utilizan para designar el destino de los mensajes de salida, por ejemplo a una aplicación servidora, una cola JMS, o una base de datos.

Mirth Connect se basa en una serie de bibliotecas que permiten el análisis efectivo de mensajes HL7, a la vez que provee un framework completo para la interacción con distintos dispositivos médicos. Mirth Connect es compatible con numerosos protocolos de transferencia utilizados en toda la industria de la salud: HTTP, Base de datos, PDF, TCP / IP, Servicios Web (SOAP), DICOM, entre otros. También es compatible con numerosos formatos de datos y normas de datos de atención médica como HL7 V2, CDA, HL7 V3, XML, NCPDP, EDI. Mirth Connect permite el desarrollo de interfaces complejas con JavaScript, supervisar estadísticas de dicha interfaz y las conexiones, también permite la visualización y procesamiento de mensajes [4].

### 2.4 Sistema RIS/PACS

RIS (*Radiological Information System*, Sistema de Información de Radiología) es un sistema que administra los datos demográficos y clínicos de los pacientes sólo en el departamento de Radiología y sirve de enlace con los demás departamentos a nivel de sistemas. RIS posee una base de datos de la agenda de estudios que se realizaron y una base de datos de los reportes radiológicos que generaron los médicos.

PACS (*Picture Archive and Communications System*, Sistema de Comunicación y archivo de imágenes) es un sistema que almacena imágenes médicas y ayuda a los elementos que lo rodean a intercambiar dichas imágenes. Este sistema posee tres depósitos de imágenes de acuerdo a las fechas de generación de los estudios.

RIS/PACS es un sistema que tiene la funcionalidad de los dos sistemas RIS y PACS. RIS/PACS pueden ser dos sistemas enlazados por una interface o uno solo con la funcionalidad completa de los dos, todo depende del producto. El software que tiene las funciones básicas de estos dos sistemas y cumple con el estándar DICOM es un RIS/PACS [5].

### 2.5 DICOM

El protocolo de comunicación DICOM (*Digital Imaging Communications in Medicine*, Comunicación de imágenes digitales en medicina) es un estándar de información tecnológica que se utiliza virtualmente en todos los hospitales del mundo. La estructura de DICOM se diseñó en 1993, con el objetivo de asegurar la interoperabilidad de los sistemas.

Los principales motivos por los que surge DICOM son:

- Una imagen médica no tiene sentido sin los datos del paciente.
- Es más seguro introducir los datos del paciente dentro de un mismo archivo de imagen.
- Los formatos existentes como: .jpg, .gif, .tiff en esos días no alcanzaban a cubrir esas necesidades.
- Distintas imágenes del paciente tienen relación entre sí.
- Necesidad de comunicar equipos diferentes entre sí.
- Necesidad de estandarizar servicios y archivos [6].

## III. ARQUITECTURA

Para este trabajo se incluyó el registro de resultados de estudios del área de Imagenología. En la Fig. 2 se muestra un esquema de la aplicación Web que se desarrolló, así como, las tecnologías y herramientas necesarias para la comunicación con los equipos médicos del área de Imagenología.

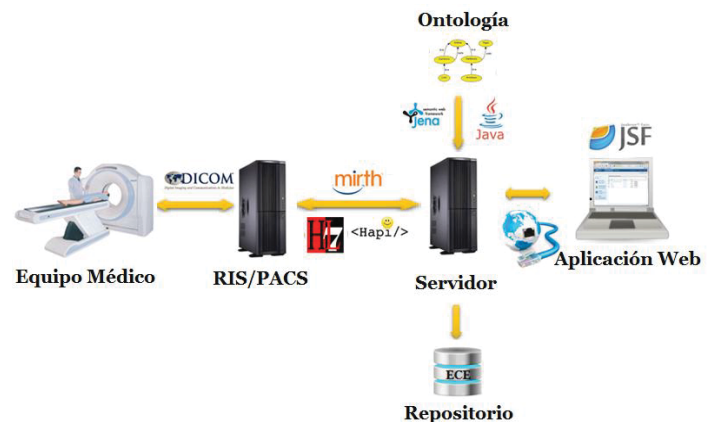


Fig. 2. Arquitectura de la aplicación Web para el registro de resultados de estudios de Imagenología

La biblioteca HAPI [7] y el motor de integración Mirth Connect (ambos basados en Java) se utilizaron para agregar capacidades de envío, interpretación y recepción de mensajes HL7 a la aplicación.

El RIS/PACS se encarga de realizar la gestión de estudios de imagen, así como de las interacciones específicas con cada equipo médico y de comunicar los resultados de dichos estudios a la aplicación Web. La aplicación Web se encarga de solicitar los estudios de Imagenología al RIS/PACS, recibir y almacenar las interpretaciones integrándolas en el expediente clínico del paciente.

Cabe mencionar que al momento de la redacción del artículo, el hospital no cuenta con un servidor RIS/PACS, necesario para hacer la comunicación con los equipos médicos, por lo tanto, se optó por desarrollar una pequeña aplicación que en conjunto con Mirth Connect representaron al RIS/PACS.

En la Fig. 3 se muestra el proceso general de envío de una solicitud de estudio al RIS/PACS y en la Fig. 4 el proceso de recepción de resultados en la aplicación Web.



Fig. 3 Envío de solicitud de estudio al RIS/PACS.



Fig. 4 Recepción de resultados de estudios en la Aplicación Web.

La aplicación Web tiene la capacidad de crear los mensajes HL7 mediante consultas a la Ontología y a las bases de datos pertinentes como se muestra en la Fig. 5.

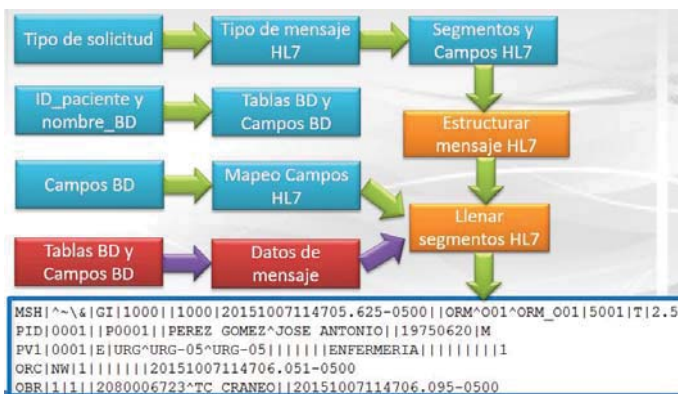


Fig. 5 Proceso de la aplicación Web para crear mensajes HL7.

Siguiendo la figura, el proceso de envío y recepción de mensajes HL7 para la aplicación, es el siguiente:

- La aplicación Web mediante la ontología es responsable de elaborar el mensaje HL7 con los datos de pacientes los cuales se enviarán como solicitudes al PACS.
- El sistema PACS comunica a la aplicación Web la aceptación o el rechazo de las pruebas solicitadas.
- Una vez finalizadas las pruebas, el PACS manda a la aplicación Web la interpretación de los estudios y la URL asociada a la dirección de la imagen. Para esto, es necesario instalar un visor de imágenes DICOM para consultar los resultados de estudios.

Durante el proceso de construcción de mensajes HL7 descrito anteriormente, se realizan tres consultas a la Ontología:

En la primera consulta, a partir del tipo de solicitud, se obtiene de la ontología el tipo de mensaje HL7 correspondiente, así mismo, de ese tipo de mensaje se obtienen sus segmentos y campos HL7, con esto se construye la estructura del mensaje HL7.

La segunda consulta a la ontología, con los parámetros: identificador del paciente y el nombre de la base de datos, permite obtener los nombres de las tablas y campos de las bases de datos en donde se aloja la información de dicho paciente.

En la tercera consulta a la ontología, se hace una búsqueda del mapeo para obtener los campos HL7 con los que están asociados cada uno de los campos de base de datos obtenidos en la segunda consulta. Además, con los nombres de las tablas y campos de la base de datos obtenidos en la segunda consulta a la ontología, se harán otras consultas a la base de datos correspondiente para obtener la información del paciente.

Una vez que se construya la estructura del mensaje, se tenga el mapeo de los campos de base de datos con los campos HL7, y se obtenga la información del paciente, entonces se llenan los segmentos del mensaje HL7.

#### IV. DESARROLLO

La arquitectura mostrada se basó fuertemente en el proceso de envío y recepción de mensajes HL7, por lo que el primer paso en el desarrollo de la aplicación fue identificar las características de tales mensajes así como los elementos para obtenerlos, en este caso con la herramienta Mirth Connect.

En la Fig. 6 se muestra el código para llenar el segmento PID de un mensaje HL7 que es el que contiene los datos de identificación del paciente. Primero, se buscan los datos de la conexión al repositorio específico en la ontología (línea 4), después se realiza una búsqueda para obtener el mapeo de los campos de la base de datos con los campos del mensaje HL7 (línea 5). Posteriormente, de acuerdo al identificador del paciente que se ingresó (línea 3) se obtiene la información del mismo (línea 6). Al final, se llena cada uno de los campos del mensaje HL7 con los datos obtenidos del repositorio (líneas 10 a 16).

```

1  PID pid = message.getPATIENT().getPID();
2  try {
3    oPaciente.setIdPaciente(idPaciente);
4    oPaciente.buscarDatosConexion();
5    oPaciente.buscarMapeo();
6    oPaciente.buscar();
7  }catch (Exception ex) {
8    Logger.getLogger(ConsultaJB.class.getName()).log(Level.SEVERE, null, ex);
9  }
10 pid.getSetIDPID().setValue(oPaciente.getIDPID());
11 pid.getPatientIdentifierList(0).getIDNumber().setValue(oPaciente.getIdPaciente());
12 pid.getPatientName(0).getFamilyName().setSurname(oPaciente.getApe
13 Paterno()+oPaciente.getApeMaterno());
14 pid.getPatientName(0).getGivenName().setValue(oPaciente.getNombre());
15 pid.getDateTimeOfBirth().getTime().setValue(oPaciente.getFechaNac());
16 pid.getAdministrativeSex().setValue(oPaciente.getSexo());
    
```

Fig. 6 Código para llenar el segmento PID

En la Fig. 7 se presenta parte del código para realizar la búsqueda de información del paciente en la base de datos, los nombres de los campos de esa consulta se obtuvieron de la ontología a través de una consulta SPARQL.

```

1  if (this.getIdPaciente().equals("")) {
2      throw new Exception("Error de programación, faltan datos");
3  }
4  }
5  else {
6      sQuery = "SELECT t1." + this.getCamposBD().get(5) + ", t1." +
7      this.getCamposBD().get(2) +
8      ", t1." + this.getCamposBD().get(4) + ", t1." + this.getCamposBD().get(0) +
9      ", t1." + this.getCamposBD().get(1) + ", t1." + this.getCamposBD().get(6) +
10     ", t1." + this.getCamposBD().get(3) +
11     " FROM " + this.getTabla() + " t1" +
12     " WHERE t1." + this.getCamposBD().get(2) + " = " + this.getIdPaciente() + """;
13     if (getoAD() == null){
14         setoAD(new AccesoDatos());
15         getoAD().setUrl("jdbc:postgresql://localhost/" + this.getBD());
16         if (getoAD().conectar()){
17             rst = this.getoAD().ejecutarConsulta(sQuery);
18             getoAD().desconectar();
19         }
20         setoAD(null);
21     }
22     else{
23         rst = this.getoAD().ejecutarConsulta(sQuery);
24     }
25 }

```

Fig. 7 Código para buscar información del paciente en el repositorio.

Por otro lado, en la Fig. 8 se muestra el código para el envío del mensaje HL7 generado por la aplicación a la herramienta Mirth Connect. El puerto escuchador de Mirth Connect que se utilizó para este proyecto fue el 6661 y la emisión del mensaje se hizo a través del protocolo TCP.

```

1  int port = 6661;
2  boolean useTls = false;
3  HapiContext context = new DefaultHapiContext();
4  Parser p = context.getPipeParser();
5  Message msg = p.parse(mensaje);
6  Connection = context.newClient("localhost", port, useTls);
7  Initiator initiator = connection.getInitiator();
8  Message response = initiator.sendAndReceive(msg);
9  String responseString = p.encode(response);
10 connection.close();
11 setResultado("La solicitud ha sido enviada y aceptada correctamente.");

```

Fig. 8 Código para buscar información del paciente en el repositorio.

## V. RESULTADOS

A continuación se muestran los resultados obtenidos de este trabajo. Los canales de transmisión de mensajes creados en Mirth Connect, y la aplicación Web poseen la capacidad de adaptarse fácilmente a las necesidades de comunicación en el momento que el hospital adquiera el servidor de imagen (RIS/PACS), por lo que los cambios en la configuración serán mínimos.

### 5.1 Creación de canales en Mirth Connect

Para que las solicitudes de estudios emitidas por la aplicación Web se reciban en la herramienta Mirth Connect y ésta devuelva la respuesta a dicha petición de estudios, se crearon dos canales de comunicación respectivamente como se muestra en la tabla 1.

Tabla 1. Canales en Mirth Connect para envío y recepción de los mensajes HL7

Nombre del Canal	Origen	Destino
HL7 Test Response	Aplicación Web	Mirth Connect
HL7 Test Result	Aplicación PACS	Aplicación Web

A través de estos canales se envían mensajes HL7 correspondientes a las solicitudes de estudios y a los resultados de los mismos. La comunicación es a través del protocolo TCP. Tanto la aplicación PACS como la aplicación Web emiten y reciben los mensajes en la versión 2.5 del estándar HL7.

### 5.2 Escenarios de integración

Los escenarios de integración identificando el ámbito de la solicitud son:

- Escenario A: Solicitud de estudios de Auxiliares de Diagnóstico.  
Los principales datos a enviar son los siguientes:
  - Datos de identificación del paciente.
  - Datos de la solicitud.
  - Datos de las pruebas estudios solicitados.
  - Datos necesarios para realizar los estudios.
- Escenario B: Recepción de resultados  
Al recibir los resultados de los estudios la aplicación Web:
  - Registra los resultados en el sistema.
  - Muestra las interpretaciones de resultados enviadas por la aplicación PACS.

### 5.3 Mensajes de integración

Solicitud de estudios de Auxiliares de Diagnóstico de Imagenología

- La aplicación Web envía a través de Mirth Connect un mensaje HL7 de tipo ORM^O01 realizando la petición de los estudios de Imagenología. El objetivo de este mensaje es comunicarle a la aplicación PACS que proceda a realizar el estudio. En la cabecera del mensaje de solicitud como aplicación origen aparece el código de la aplicación Web y como entidad receptora el código del sistema PACS.

Recepción de resultados de estudios de Auxiliares de Diagnóstico de Imagenología

- Una vez finalizadas los estudios, la aplicación PACS manda un mensaje HL7 de tipo ORU^R01 enviando la interpretación y la URL del mismo. Mirth Connect se encarga de transmitir este mensaje a la aplicación Web.

Cabe mencionar que se emplea el mismo tipo de mensaje ORM^O01 para todos los mensajes de solicitudes de estudios de Imagenología y la comunicación con las diferentes modalidades (equipos médicos). Así mismo, se utiliza el mismo tipo de mensaje ORU^R01 para todos los mensajes de resultados de estudios de Imagenología.

## 5.4 Catálogo de mensajes

En la tabla 2 se presentan los mensajes a utilizar en los escenarios de esta integración.

Tabla 2. Mensajes para los escenarios de integración identificados

Solicitud de estudios			
Tipo de Mensaje	Origen	Destino	Comentarios
ORM^O01	Aplicación Web	Aplicación PACS	Solicitud de estudios de Imagenología
Resultados de estudios			
Tipo de Mensaje	Origen	Destino	Comentarios
ORU^R01	PACS	Aplicación Web	Envío de interpretación y URL de resultados de estudios de Imagenología

## 5.5 Estructura de los segmentos

Aunque existe una gran variedad de segmentos para estos tipos de mensajes se hizo una selección de los que contienen la información más relevante para realizar el intercambio de información. Los segmentos obligatorios para cada tipo de mensaje HL7 de los escenarios de integración que se eligieron se muestran en la tabla 3.

Tabla 3. Segmentos para mensajes HL7 de los escenarios de integración identificados

Mensaje	Segmentos
ORM^O01	MSH, PID, PV1, ORC, OBR
ORU^R01	MSH, PID, PV1, ORC, OBR, OBX

Para el envío de las solicitudes de estudios de Imagenología se requiere de un mensaje ORM^O01 que es enviado desde la aplicación Web a la aplicación PACS y tiene la estructura que se muestra en la tabla 4.

Tabla 4. Segmentos del Mensaje ORM^O01

Segmento	
MSH	Cabecera del mensaje. Solo uno por mensaje.
PID	Identificación del paciente. Solo uno por mensaje.
PV1	Información acerca del episodio.
ORC	Datos generales de la solicitud.
OBR	Detalle de solicitud.

Para el envío de la interpretación de los resultados de estudios se requiere de un mensaje ORU^R01 que será enviado desde la Aplicación PACS a la Aplicación Web y tendrá los mismos segmentos que el mensaje ORM^O01 mostrados anteriormente más el segmento OBX en donde se mandan los resultados del estudio. Si el OBR indica una prueba simple, sólo tendrá un OBX dependiente de ella.

## 5.6 Llenado de los segmentos

La aplicación Web mediante consultas a la ontología tiene la capacidad de estructurar cada mensaje correspondiente a

solicitudes de estudios, además de obtener y rellenar la información correspondiente a cada segmento.

Esta aplicación Web requiere interacción con la base de datos del hospital que la utilice para la obtención de los datos con los que se llenarán los segmentos de los mensajes. Por lo anterior, se agregaron capacidades para que dicha aplicación realice conexiones a diferentes repositorios.

Antes de utilizar la aplicación Web en un nuevo hospital será necesario agregar una nueva configuración a la ontología: Nombre de la base de datos, y nombre de las tablas y campos en donde se encuentra los datos de pacientes y resultados de estudios de Diagnóstico. Además se agregará a la aplicación el driver correspondiente para conectarse a la base de datos del hospital correspondiente.

La aplicación Web que se desarrolló en este trabajo es capaz de funcionar con diferentes equipos de Imagenología y con diferentes Sistemas de ECE, para esto, basta con actualizar la ontología y hacer cambios mínimos en el código fuente de la aplicación.

## 5.7 Ejemplo de mensaje HL7 de resultados de estudios

La aplicación Web permite la construcción automática de mensajes HL7, a continuación se presenta un ejemplo de mensaje HL7 V2.x generado al solicitar un estudio de Rayos X de cráneo, la interpretación del resultado de dicho estudio tiene la siguiente estructura:

```
MSH|^~\&|MIRTH|RADIOLOGY|1000|1000|20151127||ORU^R01|1|T|2.5
PID|3||P0003||VelascoTorres^Genaro
PV1|0001|E|URG^URG-05^URG-05|||||ENFERMERIA|||||1
ORC|RE|204|||||20151127103235.453-0600
OBR|1|1||2080006723^RAYOS X DE CRANEO||20151127103235.454-0600
OBX|1|CE||18^18.dcm
```

## 5.8 Implementación de la aplicación Web

En este trabajo surgió la limitante de no contar con un Servidor RIS/PACS para obtener los resultados de estudios directamente de los equipos médicos, por lo tanto, las pruebas se realizaron con un emulador y se tuvieron resultados aceptables. Sin embargo, se planea probar la aplicación en un escenario real en donde se seguirá utilizando la herramienta Mirth Connect como intermediario entre la aplicación y el RIS/PACS.

La interfaz gráfica de usuario de la aplicación es intuitiva y tiene un menú simple. Cuando el usuario inicia sesión se presenta un menú que consta de dos secciones principalmente: solicitudes y resultados.

En la Fig. 9 se muestra el formulario de la página de solicitudes para ingresar los datos de la solicitud y enviarla al PACS (representado por el emulador) el cual se encargará de la obtención del resultado de dicho estudio.

**Datos de la solicitud**

**Ingresa el ID del paciente**

**Selecciona el tipo de estudio**

Rayos X (RX)

**Selecciona la modalidad**

Cráneo

Fig. 9 Formulario de la solicitud de estudios

La página de resultados permite la consulta de los mismos mediante diferentes filtros de búsqueda. En la parte superior de la misma se localiza un botón para actualizar los estudios que se recibieron del PACS. Además, en cada uno de los resultados se muestra un botón “Descargar” para obtener el archivo en formato DICOM correspondiente al estudio. Para abrir el archivo de resultados es necesario contar con un visor DICOM, por ejemplo RadiAnt-DICOM como se muestra en la Fig. 10.

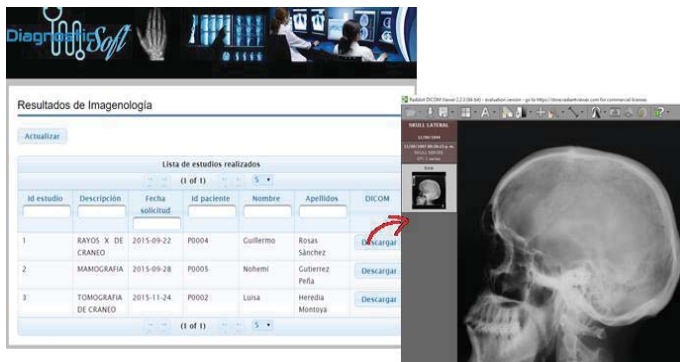


Fig. 10. Resultado de estudio de Imagenología

## VI. TRABAJO A FUTURO

Como el estándar HL7 ofrece una diversidad de mensajería en informática médica, se seguirá trabajando para enriquecer los conceptos en la ontología para que soporte la creación de otros tipos de mensajes. Aunque la aplicación Web desarrollada está habilitada para crear mensajes y comunicarse con equipos del Área de Imagenología, se le agregarán capacidades para la construcción de mensajes del área de Laboratorio, y se crearán los canales de comunicación necesarios para interactuar con un Sistema Integral LIS el cual gestionará los resultados de estudios de Laboratorio a los equipos médicos de esta área.

Por otro lado, se propone agregar a la aplicación Web presentada en este proyecto una sección para configurar los datos necesarios para que funcione en diferentes hospitales. Se plantea que el usuario indique los parámetros de comunicación (como el nombre del repositorio, nombre del PACS, driver de

la conexión, entre otros) requeridos entre dicha aplicación y el Sistema de ECE y PACS correspondientes, de modo que la ontología se actualice de forma automática con los nuevos datos de otros hospitales desde la aplicación Web.

## VII. CONCLUSIONES

En este trabajo se destaca la importancia del uso de una aplicación Web para complementar un sistema de ECE con los resultados de estudios, mediante una ontología y siguiendo las especificaciones del estándar HL7 y la Norma Oficial Mexicana en su desarrollo. Dicha aplicación tiene las siguientes ventajas:

- Facilita la obtención de los datos que requiere cada segmento de un mensaje HL7.
- Ahorra tiempo en la creación de los mensajes HL7.
- Permite establecer la comunicación entre diferentes equipos y sistemas de ECE independientemente del fabricante.
- Brinda una interfaz gráfica fácil de usar.

## AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACyT), por el apoyo económico otorgado para la realización de la investigación.

A la empresa internacional everis México, S. de R.L. de C.V por las facilidades otorgadas para la realización de una estancia académica en sus instalaciones que permitió enriquecer el presente trabajo.

A la empresa Servicios Profesionales en Imagen Médica (SPIM) por el aporte de conocimiento a este trabajo mediante los cursos “Tomografía Axial Computarizada” y “Comunicación de equipos médicos radiológicos con Sistemas de Información”

## REFERENCIAS

- [1] NORMA Oficial Mexicana NOM-024-SSA3-2012, Sistemas de expediente clínico electrónico. Disponible en <http://www.dof.gob.mx/normasOficiales/4151/salud/salud.htm>
- [2] Laboratories Exakta. Importancia de los Métodos Auxiliares de Diagnóstico en la Salud. Disponible en: <http://exakta.goplek.com/contenido/463/Importancia-de-los-M%C3%A9todos-Auxiliares-de-Diagn%C3%B3stico-en-la-Salud.html>
- [3] Sitio oficial de HL7. Disponible en: <http://www.hl7.org>
- [4] Sitio Oficial Mirth. Disponible en: <http://www.mirthcorp.com/products/mirth-connect>
- [5] Gómez Otero Alfredo. Comunicación de equipos médicos radiológicos con Sistemas de Información. Material de taller online, 2015.
- [6] Sitio Oficial RadiAnt DICOM. Disponible en: <http://radiantviewer.com/>
- [7] Sitio Oficial HAPI. Disponible en: <http://hl7api.sourceforge.net/>

# Sistema de Información para el Control de Acceso a Datos Personales

Olimpia Saucedo Estrada, Alejandra Giselle Sánchez  
Ortiz, Alexia Beltrán Ortega, Jorge Ignacio Vázquez  
Hernández, Gina Gallegos-García  
Escuela Superior de Ingeniería Mecánica y Eléctrica.  
Unidad Culhuacan. Instituto Politécnico Nacional  
Ciudad de México. México.  
ggallegosg@ipn.mx

Gualberto Aguilar Torres  
Comision Nacional de Seguridad. Secretaria de  
Gobernacion.  
Ciudad de México. México  
gualberto.aguilar@cns.gob.mx

**Resumen— Actualmente los Sistemas de Información para el Control de Acceso a Datos Personales validan las credenciales de una entidad, a través de técnicas que hacen uso de un solo factor de autenticación, entre los que destacan: la contraseña, el certificado digital y algún medio biométrico, lo cual trae como consecuencia que existan vulnerabilidades en dichos Sistemas pues no se tiene la certeza de que el cliente y el servidor sean legítimos. Con base en ello, en el presente documento se presenta el desarrollo de un Sistema de Información para el Control de Acceso a Datos Personales el cual utiliza dos factores de autenticación para brindar una mayor seguridad al Sistema de Información para el Control de Acceso a Datos Personales.**

**Palabras Clave—Control de Acceso; Certificado Digital; Datos Personales; Firma digital; Función Hash; Protocolo desafío-respuesta; Sistema de Información.**

## I. INTRODUCCIÓN

La autenticación de los usuarios, es un servicio de seguridad encargado de controlar el acceso dentro de un sistema de información orientado al tratamiento y administración de datos personales.

El procedimiento para la autenticación permite validar que los usuario sean quienes dicen ser y no terceras entidades que hayan obtenido las credenciales sin ser los usuarios auténticos.

Para llevar a cabo esta tarea de autenticación, se requiere hacer uso de mecanismos y protocolos de seguridad, que se ejecuten dentro de un sistema de información, de tal forma que se puedan intercambiar los identificadores del usuario y un servidor.

En la actualidad, los controles de seguridad que utilizan estos sistemas de información, no implementan protocolos de autenticación mutua, por lo que las entidades participantes, solo autentican a una de ellas, por medio de un factor único. Lo anterior, da pie a la existencia de vulnerabilidades en los sistemas de información enfocados al tratamiento y administración de datos personales, dado que no se está seguro de que el cliente y el servidor sean quienes dicen, surgiendo la necesidad de autenticar ambos lados.

Con base en lo anterior, en este artículo se presenta el desarrollo de un sistema de información, el cual a través de un mecanismo de autenticación, permite el acceso a datos personales, por medio de dos factores de autenticación, de tal forma que el sistema de información en el cual se encuentra implementado dicho control, autentica mutuamente al cliente como al servidor. Es importante destacar que este sistema considera las recomendaciones que dicta la Ley Federal de Protección de Datos Personales en Posesión de Particulares (LFPDPPP) en referencia a este tipo de información sensible [1].

## II. TRABAJOS RELACIONADOS

Algunas de las propuestas que se tomaron como base para el desarrollo del Sistema de Información para el Control de Acceso a datos Personales se citan a continuación.

En el 2007, en [2] se presenta el desarrollo de un sistema de acceso a servicios escolares con dos factores de autenticación, una contraseña y un protocolo de seguridad alojado en tarjetas inteligentes, dichas tarjetas almacenaban los datos personales. Para que el usuario se autentique y obtenga acceso a la información es necesario que ingrese su contraseña.

En 2012 en [3] los autores procedentes de la Pontificia Universidad Católica del Perú, presentan un sistema de acceso a diversos servicios escolares. Esto, haciendo uso de una contraseña única para todas sus plataformas, como único factor de autenticación.

En el mismo año, en [4] se discuten tres factores de autenticación para que un usuario autorizado pueda acceder a un sistema de información: el primero es el nombre de usuario y contraseña, el segundo es una contraseña dinámica y el tercero un certificado digital. Se concluye en el artículo que el certificado digital da mayor confidencialidad en un usuario autorizado.

En 2014 en [5] se propuso un sistema que permitía realizar consultas seguras en una base de datos mediante certificados dinámicos, en el cual se genera un número aleatorio y se le agrega un número entero muy largo. Este es cifrado cuando se



hace la petición. También se compara el certificado proporcionado con el que se tiene almacenado.

Como se puede ver, los sistemas de acceso a datos personales, aún tienen ciertas desventajas que dejan abierta la necesidad de autenticar mutuamente tanto al cliente como al servidor, pues al no contar con una autenticación mutua existen vulnerabilidades debido a que no se está seguro de que el cliente y el servidor sean quienes dicen ser.

### III. REQUISITOS DE LAS ETAPAS DEL SISTEMA DE INFORMACIÓN PARA EL CONTROL DE ACCESO A DATOS PERSONALES

Con base en las desventajas que presentan los trabajos relacionados, en esta sección se listan los requisitos que debe tener el Sistema de Información para el Control de Acceso a datos Personales.

El sistema de información debe tener dos etapas, el registro de los usuarios y el acceso al sistema de información. Estas etapas se proponen para implementar dentro del mismo, el mecanismo de autenticación que permite el acceso a datos personales, por medio de dos factores de autenticación. Para ello, estas dos etapas constan de los siguientes requisitos:

1. Se requiere realizar el registro de los nuevos usuarios, añadiendo sus datos personales.
2. Se debe pedir al usuario que agregue una contraseña.
3. Se pide el acceso al sistema para consulta de datos.
4. En caso de que el usuario esté previamente registrado, solo pedirá el acceso para poder obtener sus datos personales.
5. El sistema debe pedir al usuario su contraseña y nombre para dar el acceso a los datos.
6. Se requiere insertar el dispositivo USB para ingresar a los datos.

### IV. PRIMITIVAS CRIPTOGRÁFICAS UTILIZADAS EN LAS FASES DEL SISTEMA DE INFORMACIÓN PARA EL CONTROL DE ACCESO A DATOS PERSONALES

Debido a que el Sistema de Información para el Control de Acceso a Datos Personales debe tener un manejo seguro de los datos que poseen, se requieren del uso de aquellas primitivas criptográficas que garanticen los servicios de seguridad necesarios, entre los que destacan las funciones hash y las firmas digitales.

#### A. Función Hash

Una función hash es una función conmutable mediante un algoritmo que acepta un mensaje ( $m$ ) de entrada y genera uno de salida  $H(m)$ , llamado resumen del mensaje [6].

Las funciones hash más habituales en el campo de la criptografía provienen de la familia SHA [7]. En la Figura 1 se ilustra un diagrama de una función hash en donde se puede observar un mensaje que entra a la función  $H$  y a la salida se tiene un resumen de este mensaje, el cual es único para esa entrada de mensaje debido a que no puede ser repetido por otra entrada diferente.

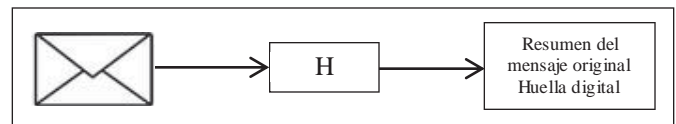


Figura 1. Diagrama general de una función hash

SHA-2 consiste en un conjunto de cuatro funciones hash de 224, 256, 384 o 512 bits. SHA 256 usa palabras de tamaño de 32 bits. La función SHA-2 se implementa en el Sistema de Control en la etapa de registro, de tal forma que se pueda obtener una huella digital de todos los campos del certificado emitido por el usuario, haciendo que dicho certificado sea el segundo factor de autenticación.

#### B. Firma Digital

La firma digital es una cadena de datos, que asocia un mensaje de manera digital, con la entidad que origina el mensaje. Esta primitiva criptográfica, es fundamental en la autenticación y no repudio [7].

El proceso de firma y el de verificación se utiliza en las etapas de registro y autenticación. En la etapa de registro, se obtiene el valor hash de los datos personales del usuario. Este valor hash resultante es el que se firma con la llave privada del firmante y se concatena a los datos personales originales. En la etapa de acceso, el sistema separará el mensaje y la firma. La firma se verificará con la llave pública del firmante. Si la verificación es verdadera, el usuario es auténtico y puede acceder al sistema.

### V. SEGURIDAD EN EL SISTEMA DE INFORMACIÓN PARA CONTROL DE ACCESO A DATOS PERSONALES

Las primitivas criptográficas que se utilizan dentro del Sistema de Información para el Control de Acceso a Datos Personales, requieren del uso de llaves criptográficas. La forma en cómo se utilizan dichas llaves, se menciona a continuación.

#### A. Certificado Digital X.509 v3

El enlace entre un usuario y su llave pública se realiza cuando una Autoridad Certificadora (AC) firma cada certificado. Un certificado digital incluye información del propietario y la AC que lo respalda [8].

Al crear el certificado, la información contenida en él, es firmada digitalmente por la AC. La firma digital de la AC sirve para detectar que el contenido del certificado no ha sido alterado y para autenticar la llave del usuario [7].

El estándar internacionalmente aceptado para certificados digitales es el X.509, en su versión 3. Esta especificación agrega campos denominados extensiones, donde se pueden definir características de los certificados inherentes a cada entidad. En la Figura 2 se muestra la estructura de un certificado digital X.509 versión 3.

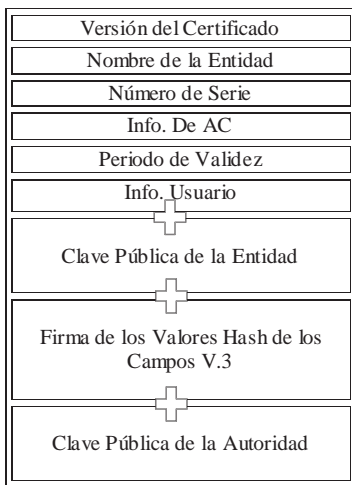


Figura 2. Estructura de certificado digital

Este certificado se genera en la etapa de registro al momento en el que el usuario proporciona sus datos personales. Posteriormente, se almacena en un dispositivo USB.

### B. Protocolo Desafío-Respuesta basado en Firmas Digitales

La autenticación de desafío-respuesta es un protocolo en el cual la parte verificadora presenta una pregunta o desafío y la parte a ser verificada debe dar una respuesta válida para ser autenticado.

En este protocolo de seguridad se comprueba la información de la cuenta del usuario mediante un intercambio cifrado de mensajes en los que no se transmite la contraseña real. La idea de los protocolos criptográficos desafío-respuesta es que una entidad “provee” su identidad a otra entidad demostrando conocimiento de un secreto que está asociado con dicha entidad, sin revelar el secreto al verificador durante el protocolo. Esto se realiza proporcionando una respuesta a un desafío variante en el tiempo. El desafío es típicamente un número pseudoaleatorio elegido por una entidad al inicio del protocolo. Si la línea de comunicación es intervenida, la respuesta de una ejecución del protocolo de identificación no proporciona a un adversario información útil para una subsiguiente identificación. Una vez ejecutado este protocolo, el sistema es capaz de permitir el acceso a él o si fuera el caso y no se haya comprobado la identidad [7].

## VI. DISEÑO DEL SISTEMA DE INFORMACIÓN PARA EL CONTROL DE ACCESO A DATOS PERSONALES

El sistema de información propuesto contiene dos etapas importantes, la etapa de registro y la etapa de acceso.

En la etapa de registro se obtienen los dos factores de autenticación, en este caso la contraseña del usuario y el certificado digital. En la etapa de acceso se autentica al usuario al momento en el que él acceda a sus datos personales.

Para el diseño del sistema de información se generaron diagramas de casos de uso y flujo de eventos de cada una de las fases. Estos se describen a continuación.

### A. Diagrama de Casos de Uso del Sistema de Información para el Control de Acceso a datos Personales

Los casos de uso son herramientas de modelización de requisitos funcionales, que preceden en tiempo y exceden en alcance a UML.

Estos diagramas se utilizan para describir modelos de sistemas orientados a objetos. Se basan en escenarios para la obtención de requerimientos e identifican el tipo de interacción y los actores involucrados [9].

La Figura 3 muestra el diagrama de casos de uso del funcionamiento general para el SICADP, en donde se ve involucrado el actor llamado usuario no registrado, quien interactúa con el caso de uso de inicio y de registro. Como segundo actor, se tiene al usuario registrado, que interactúa con los casos de uso de inicio y acceso.

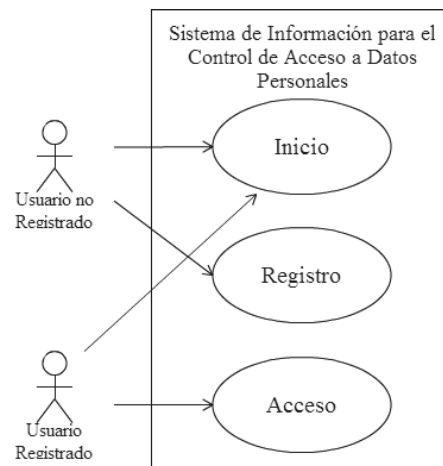


Figura 3. Diagrama general de casos de uso del sistema de información para el control de acceso a datos personales

### B. Diagrama de caso de uso del Inicio

Este caso de uso permite la interacción entre el SICADP y los actores registrados y no registrados. Se da a elegir entre registrarse como nuevo usuario o acceder al sistema si ya cuenta con un registro, como se muestra en la Figura 4.

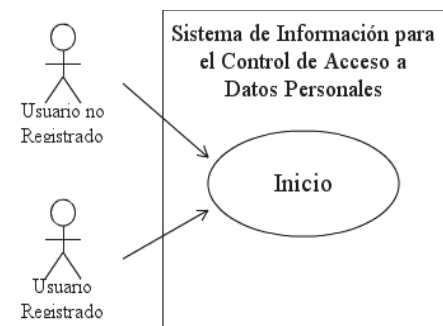


Figura 4. Diagrama de casos de uso para el inicio del sistema de información

C. Etapa de Registro

En esta fase se realiza el almacenamiento de los datos personales de cada usuario que desee registrarse y se genera un certificado digital que contiene información del usuario y su clave pública. Dicho certificado es almacenado en una unidad USB para su posterior uso en el caso de uso de Acceso.

D.1. Diagrama de caso de Uso para la etapa de Registro

En la Figura 5 se muestra el caso de uso para la etapa de registro. En esta, el actor es un usuario no registrado el cual solicita registro de sus datos personales, al término de esta etapa será creado el certificado digital para cada usuario nuevo.

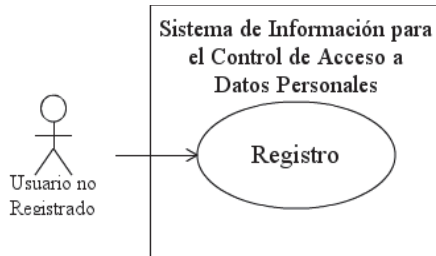


Figura 4. Diagrama de Caso de Uso de la etapa de Registro

D. Etapa de Acceso

Esta etapa se encarga de llevar a cabo el proceso de autenticación mutua por medio del protocolo desafío-respuesta entre el usuario previamente registrado y el Servidor. Solamente cuando estos dos actores se autentican se permite el acceso a la información del usuario, de caso contrario se niega el acceso. Para que el servidor de acceso a la información, el usuario debe proporcionar sus factores de autenticación para validar su identidad.

E.1. Diagrama de caso de Uso para la etapa de Acceso

En la Figura 5 se muestra el diagrama de casos de uso para la etapa de acceso, en la que un usuario registrado solicita el acceso a sus datos proporcionando sus factores de autenticación para validar su identidad y ser acceder a sus datos como un usuario auténtico.

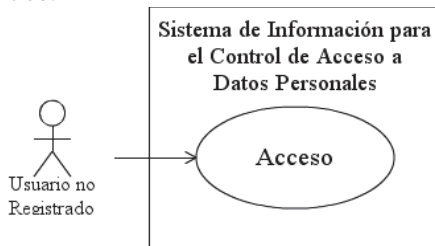


Figura 5. Diagrama de Caso de Uso de la etapa de Acceso

Posteriormente en la Figura 6 se puede observar el flujo de la información entre en usuario y el Sistema.

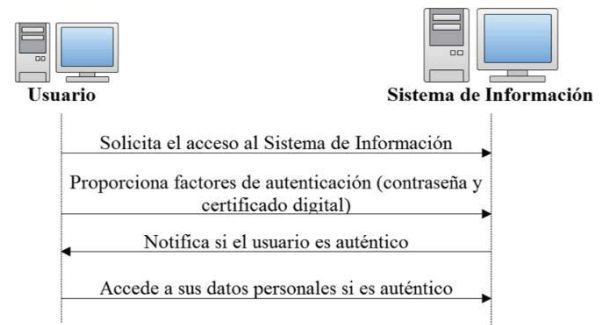


Figura 6. Flujo de Información de la etapa de Acceso

VII. PRUEBAS EFECTUADAS AL SISTEMA DE INFORMACIÓN PARA EL CONTROL DE ACCESO A DATOS PERSONALES

A continuación se muestran las pruebas a las que fue sometido el Sistema de Información para el Control de Acceso a Datos Personales y los resultados que se obtuvieron de ellas. Se realizaron cuatro tipos de pruebas, en las cuales se analizó si Sistema de Información cumple con los requisitos y el objetivo ante la problemática presentada.

A. Pruebas de Defecto- Análisis Dinámico (Caja Negra)

Estas pruebas se llevaron a cabo para verificar que el Sistema de Información cumpla con las funciones para las cuales fue creado, por tanto deberá cumplir con una etapa de Registro y posteriormente una etapa de Acceso.

A.1. Fase de Registro

Primeramente se muestra la interfaz gráfica para el Usuario, en donde se puede seleccionar la opción de Nuevo Usuario para registrar sus datos personales y se realiza una confirmación a esta petición.

Al confirmar la petición se despliega el formulario que deberá ser llenado por el Usuario proporcionando sus datos personales y una contraseña. Esto se puede ver en la Figura 7.

Figura 7. Fase de Registro: Formulario

Una vez realizado esto se oprime el botón “Registrar Nuevo Usuario”, seguido a esta acción, se introduce la ubicación del dispositivo USB como se muestra en la Figura 8.

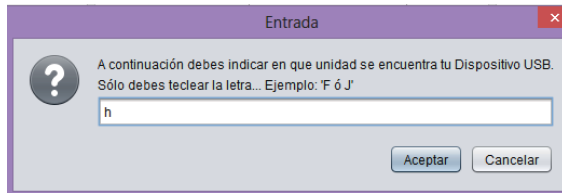


Figura 8. Ubicación de dispositivo USB

Después se valida que la contraseña fue exitosa, como se muestra en la Figura 9.

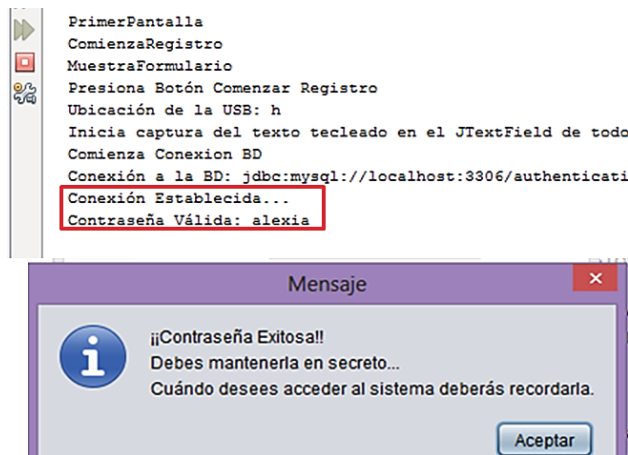


Figura 9. Validación de Contraseña mostrada en consola

Posteriormente se genera el Certificado Digital, como se puede observar en la Figura 10.

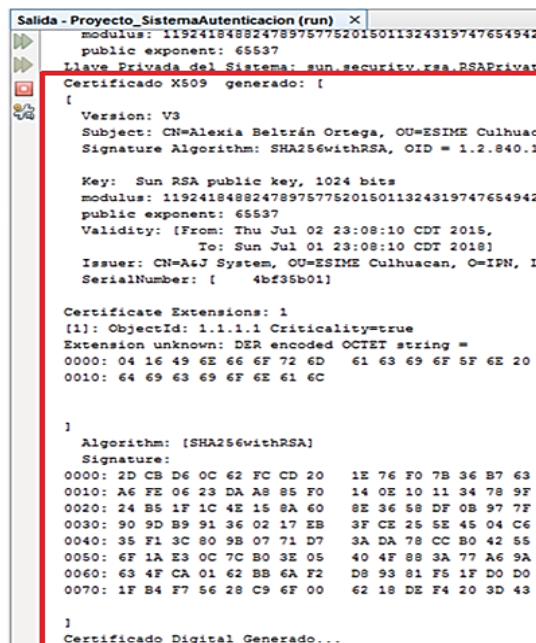


Figura 10. Creación del certificado digital en consola

Por último éste es almacenado en el dispositivo. De esta manera es cómo termina la etapa de Registro como se observa en la Figura 11.

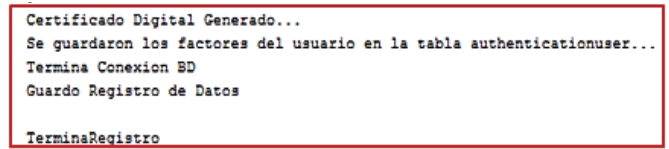


Figura 11. Término del Registro mostrado en consola

### A.2. Fase de Acceso

Para las pruebas realizadas en ésta fase, se tiene como primera ventana las instrucciones y requisitos para lograr acceder al Sistema de Información. El usuario proporciona su contraseña y la ubicación del dispositivo USB en donde se encuentra almacenado su Certificado Digital que deberá ser extraído como se muestra en la Figura 12.

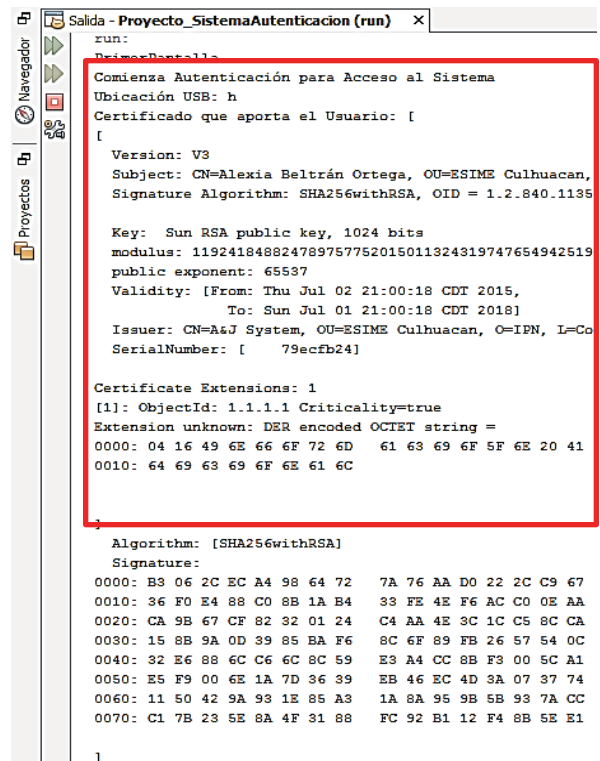


Figura 12. Extracción del certificado digital mostrado en consola

Una vez que el certificado es extraído, se solicita la llave privada al Usuario para validar el Certificado y así mismo se valida el protocolo de autenticación mutua, como se observa en la Figura 13.

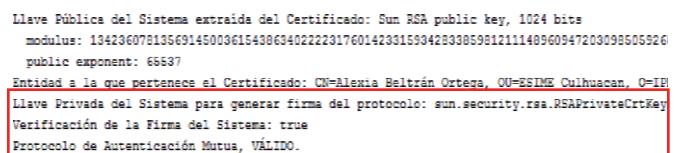


Figura 13. Llave privada extraída mostrada en consola

A continuación en la Figura 14 se realiza la búsqueda de la contraseña concatenada con el número de serie del certificado en la base de datos para corroborar que el Usuario forme parte del SICADP y se comienza el protocolo de autenticación mutua para así poder acceder al Sistema de Información.

```

]
Llave Pública del Sistema extraída del Certificado: Sun |
modulus: 13423607813569145003615438634022231760142331|
public exponent: 65537
Entidad a la que pertenece el Certificado: CN=Alexia Bel
Llave Privada del Sistema para generar firma del protoco
Verificación de la Firma del Sistema: true
Protocolo de Autenticación Mutua, VÁLIDO.
Comienza Conexión BD
Conexión a la BD: jdbc:mysql://localhost:3306/authenticac
Conexión Establecida...
Valor Autenticación: a51c31950a4721e65505f0b7ca86235f782c
Consulta de la Tabla authenticationuser: SELECT infoUsua
Busqueda del valor concatenado en la bd...
Resultado de la búsqueda de idUsr: 2
Termina Conexión BD
Comienza Conexión BD
Conexión a la BD: jdbc:mysql://localhost:3306/authenticac
Conexión Establecida...
Consulta de la Tabla infoUser: SELECT Nombre, ApPaterno,
Busqueda del valor infoUsuario en la bd...
Resultado Consulta: Alexia Beltrán Ortega
15 12 1993
DistritoFederal Coyoacán
Termina Conexión BD
    
```

Figura 14. Autenticación Mutua y conexión con la Base de Datos

Se procede a mostrar los datos personales del usuario como se puede ver en la figura 15.

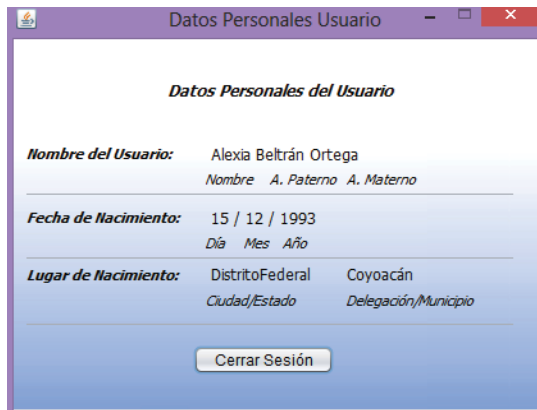


Figura 15. Fase de Acceso: "Datos Personales del Usuario"

Por último, una vez que se mostró al Usuario su Información, se termina la sesión.

### B. Caja Blanca

Estas pruebas se refieren al estricto funcionamiento del sistema. A diferencia de la prueba de defecto, en Análisis Dinámico de Caja Blanca, no se busca sólo verificar que el funcionamiento del Sistema de Información cumpla con el objetivo para el cuál fue creado, sino que verifique que el objetivo haya sido delimitado en todas y en cada una de las

etapas que se consideran para este, y que el resultado para acceder a la información del Usuario sea el más adecuado.

#### B.1.1 Fase de Registro

En estas pruebas se realizan validaciones de los campos del formulario llenado por el usuario. Debido a que son datos personales, no se pueden contener caracteres diferentes al alfabeto. Tampoco se puede insertar un número entre ellos; en el caso de la fecha de Nacimiento se considera que se indique a través del número de mes y no por el nombre el mes, un ejemplo se muestra en la Figura 16.

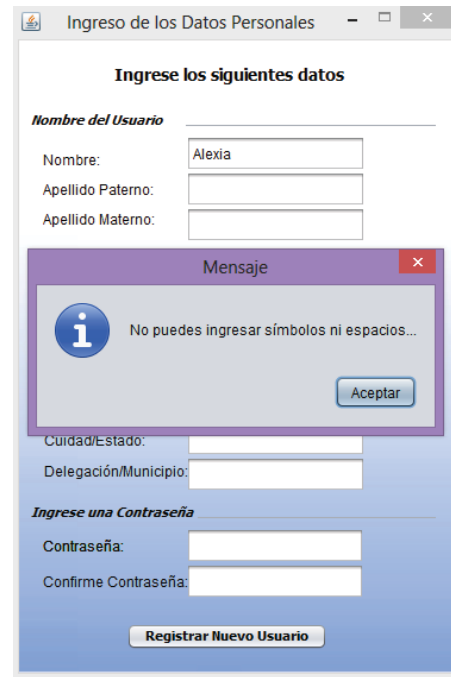


Figura 16. Errores de acceso de información.

Posteriormente se valida si el dispositivo USB contiene o no la llave pública del usuario para generar su certificado digital, la cual es fundamental para obtener su segundo factor de autenticación y que en la etapa de Acceso se pueda verificar que el Usuario es válido para el SICADP. Si el dispositivo no es válido, se genera el mensaje mostrado en la Figura 17.

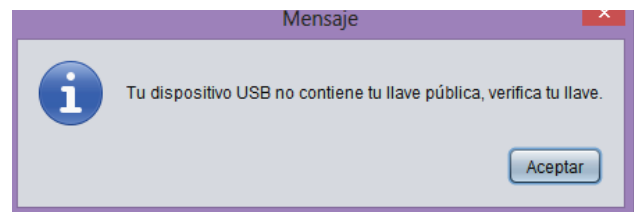


Figura 17. Validación de USB

#### B.1.2 Fase de Acceso

En estas pruebas se verifica que el certificado digital se encuentre en el dispositivo USB indicado por el usuario y que dicho dispositivo contenga la llave privada del usuario para efectuar el protocolo de autenticación mutua.

## IX. CONCLUSIONES

Se pretende adicionalmente comprobar que realmente el usuario definió la ubicación de su dispositivo USB, en caso de la carencia de alguno de estos elementos, no se podrá acceder a la información del Usuario como se muestra en la Figura 18.

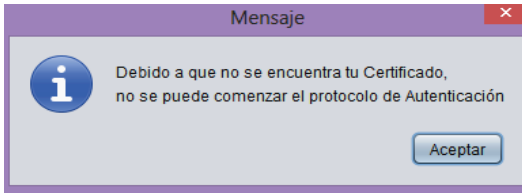


Figura. 18. Fase de Acceso: "No se encuentra Certificado Digital"

### C. Prueba Negativa

Para esta prueba se considera un escenario en dónde además de las delimitaciones que propone el Sistema de Información, se requiere saber qué pasaría si el Usuario no ingresa correctamente su primer factor de autenticación, o si no aporta su llave pública para la generación de su Certificado Digital durante su registro, otra posibilidad es que cuando desee acceder no tenga disponible en su dispositivo USB el Certificado Digital o la llave privada para comenzar el Protocolo de Autenticación Mutua.

## VIII. RESULTADOS

Los resultados que se obtuvieron al ejecutar las pruebas de defecto y pruebas de caja blanca mostraron que el Sistema de Información para el Control de Acceso a Datos Personales ofrece mejoras en comparación con los trabajos relacionados que se describieron en la Sección II.

En el trabajo presentado en [2], se mostró una propuesta de solución que consideraba dos factores de autenticación. Sin embargo, dicha propuesta no permite hacer modificaciones en el uso de algún otro protocolo de autenticación. En ese sentido, el Sistema de Información para el Control de Acceso a Datos Personales, mostró que si alguna primitiva es vulnerada, por lo que se tiene la posibilidad de mejorarla o incluso de utilizar algún otro protocolo.

En el trabajo presentado en [3] se utiliza, como factor de autenticación una única contraseña para acceder a todos los servicios de la Universidad, por lo que en caso de que un tercero no deseado obtenga dicha contraseña, se puede tener el acceso a todos los sistemas en los cuales el alumno esté registrado. El Sistema de Información para el Control de Acceso a Datos Personales además de utilizar una contraseña, utiliza un certificado digital, por lo que, para lograr el acceso a la información, se requiere aportar ambos factores.

Por último, en [5], al igual que en [2], para acceder a las bases de datos, se requiere únicamente una contraseña y un número aleatorio. Sin embargo, como se mencionó antes, los resultados obtenidos mostraron que el uso de dos factores de autenticación hace que el Sistema de Información para el Acceso a Datos Personales, lo que hace un sistema más seguro.

En la actualidad, los Sistemas de Información para el Control de Acceso a Datos Personales no cuentan con un proceso de autenticación mutua. Es por eso que surge la necesidad de crear el Sistema de Información para el Acceso a Datos Personales, ya que proporciona características y cualidades que garantizan tener una mayor seguridad, pues emplean dos factores de autenticación que son la contraseña y el certificado digital que por medio del uso de un dispositivo USB se autentica al usuario. De esta manera el Sistema de Información para el Control de Acceso a Datos Personales permite mitigar las vulnerabilidades existentes al momento de ejecutar el protocolo de autenticación mutua.

Los resultados obtenidos a través de las pruebas realizadas a nivel de fases, permitieron obtener un Sistema que solo permite acceso a usuarios autorizados por medio de los factores de autenticación. Es decir, que solo aquellos que aporten su contraseña y su correspondiente certificado digital, ligados a su registro en la base de datos, ante el Sistema, puedan acceder a sus datos personales a través de los mecanismos de control establecidos.

Con base en ello, el Sistema de Información para el Control de Acceso a Datos es una herramienta con mayor seguridad a las anteriormente desarrolladas y de gran ayuda, pues con sus dos factores de autenticación, al momento del registro y acceso de los usuarios, evita las vulnerabilidades que se puedan presentar, dando al usuario confianza y tranquilidad de que sus datos no serán vistos por terceras entidades.

## REFERENCIAS

- [1] "Ley Federal de Protección de Datos Personales en Posesión de Particulares", Diario Oficial, México D.F., 2010.
- [2] Edwin Vargas García, Vladimir Trujillo Olaya, Jaime Velasco Medina. "Diseño de un carné de identificación universitario mediante tarjetas inteligentes", Escuela de Ingeniería Eléctrica y Electrónica, Universidad del Valle, Cali, Colombia, 2007.
- [3] Normas de uso del sistema de autenticación con contraseña única, Pontificia Universidad Católica del Perú, Perú, 2012.
- [4] W. Kehe, C. Wei, and G. Yueguang, "The Research and Implementation of the Authentication Technology Based on Digital Certificates," 2012, pp. 219–222.
- [5] K. Gautam and D. Upadhyay, "Implementing dynamic certificates for securing database," in Confluence The Next Generation Information Technology Summit (Confluence), 2014 5th International Conference-, 2014, pp. 919–922.
- [6] Federal Information Processing Standards Publication, FIPS PUB 180-4, marzo 2012.
- [7] J. Alfred Menezes, "Handbook of Applied Cryptography", CRC Press, 1996.
- [8] Housley R., Ford W., Polk W., and Solo D. Internet X.509 Public Key Infrastructure, Certificate and CRL Profile. Network Working Group.
- [9] Ian Somerville, "Ingeniería de Software", Pearson Education, Madrid 2005.

# Protección de Imágenes a Color en el Dominio de la DWT Mediante Marca de Agua Invisible

Javier Molina-García<sup>1</sup>, Rogelio Reyes-Reyes, Volodymyr Ponomaryov, Clara Cruz-Ramos

Sección de Estudios de Posgrado e Investigación

Instituto Politécnico Nacional, ESIME Unidad Culhuacán

Av. Santa Ana No. 1000, Col. San Francisco Culhuacán, CP. 04430, México, D.F.

Teléfono (55) 5729-6000 Ext. 73264 Fax (55) 56562058 email: <sup>1</sup>jmolinag1001@alumno.ipn.mx

**Resumen**—Actualmente la protección de derechos de autor en archivos digitales se ve afectada debido a múltiples herramientas de procesamiento de contenido digital y al incremento de usuarios de Internet, entre los materiales afectados se encuentran imágenes, video, audio, documentos de texto, etc. Para solventar esta problemática existen diversas técnicas para la protección de derechos de autor, una de las más eficientes son las marcas de agua, las cuales pueden ser clasificadas de acuerdo a su robustez o al dominio de inserción. En el presente trabajo se propone una técnica de inserción y extracción de una marca de agua en el dominio de la Transformada Discreta de Wavelet (DWT) usando la técnica QIM-DM (Modulación por Índice de Cuantización – Modulación Dither) dentro del canal de luminancia del espacio de color YCbCr de imágenes digitales a color, programado en una plataforma de procesamiento paralelo. Los resultados obtenidos muestran que el sistema ofrece una buena calidad de la imagen marcada y es robusto ante ataques de compresión JPEG y adición de ruido impulsivo y Gaussiano. Así mismo, utilizando cómputo en paralelo el sistema requiere de tiempos de procesamiento tan pequeños que es factible su implementación para la protección de video digital.

**Palabras clave**—*Marca de Agua Invisible, Transformada Discreta de Wavelet, Modulación por Índice de Cuantización, Modulación Dither, Cómputo en Paralelo.*

## I. INTRODUCCIÓN

La violación de derechos de autor en archivos digitales es un problema que día a día se vuelve más común debido a la facilidad con la que se pueden realizar copias del contenido digital por medio de algún equipo computacional o en su defecto la Internet, en su gran mayoría estas copias realizadas al contenido digital son de forma ilegal llevando a cabo la manipulación o el uso indebido del archivo sin el consentimiento del creador del contenido digital. Para solucionar estos problemas existen diversas técnicas basadas en marcas de agua, las cuales consisten en insertar información del creador del contenido para que este tenga una forma legal de propiedad, esta información insertada puede ser invisible o visible al Sistema Visual Humano (HVS) y pueden ser utilizadas en casos legales para identificar al propietario del archivo original. A pesar de que las marcas de agua son una herramienta eficiente para contrarrestar esta problemática, un inconveniente que presentan es que en su gran mayoría poseen una alta complejidad computacional debido a todos los cálculos que realizan para insertar la marca de agua, por lo cual es necesario reducir el tiempo de procesamiento computacional a su mínima expresión, o llevar a cabo el proceso de marcado de agua de forma instantánea.

Una marca de agua digital puede definirse como un código o señal permanente, perceptible o imperceptible, incrustada en datos digitales que se puede detectar o extraer por medio de operaciones computacionales y que puede contener información acerca del propietario, de los derechos de autor, el creador, el usuario autorizado, el número de copias o reproducciones autorizadas, etc. con el fin de hacer afirmaciones acerca de los datos [1-2]. Por medio de esta afirmación es posible llevar a cabo la protección de derechos de autor. Para que una marca de agua pueda considerarse como útil y de gran eficiencia es necesario que cumpla con ciertos requisitos entre los que se encuentran: imperceptibilidad, robustez y capacidad de inserción. La imperceptibilidad se refiere a que la marca de agua puede ser perceptible o imperceptible al HVS, la robustez de la marca de agua indica la dificultad para remover la marca del archivo digital marcado, y la capacidad de inserción indica la cantidad de información que se puede ocultar por medio del algoritmo [3-5].

Las marcas de agua digitales se pueden clasificar de acuerdo a la robustez, la cual se divide en: robusta, frágil y semi-frágil; la primera debe resistir a todo tipo de ataques ya sean intencionales o no intencionales y es utilizada principalmente para la protección de derechos de autor, la segunda es utilizada para detectar cualquier tipo de modificaciones no autorizadas, y la última, es utilizada para detectar algunas modificaciones realizadas en el archivo digital original, este tipo de marcado resiste ataques no intencionales como son la compresión JPEG o la adición de ruido, pero no resiste ataques intencionales como son rotación, cambio de tamaño, etc. [3]. Por otra parte las marcas de agua se pueden clasificar de acuerdo al dominio de la inserción entre las que se encuentran la inserción en el dominio espacial en la cual los píxeles de la imagen son modificados directamente para insertar la marca de agua, y por otra parte existe la inserción en el dominio de la frecuencia [3-5] en la cual se pueden aplicar diversas transformaciones a la imagen como son la Transformada Discreta de Fourier (DFT), Transformada Discreta del Coseno (DCT), la Transformada Discreta de Wavelet (DWT) y la Descomposición en Valores Singulares (SVD) [4-5].

La Transformada Discreta de Wavelet ha tenido múltiples usos en aplicaciones de marca de agua debido a la robustez que se obtiene ante diversos ataques de procesamiento de imágenes [4-5]. Esta transformación realiza una descomposición de la imagen original en imágenes de aproximación las cuales representan las sub-bandas de frecuencia de la señal original,

estas se dividen en: una sub-banda de baja frecuencia (LL), la cual representa los datos más importantes de la señal; una sub-banda con detalles horizontales (HL); una sub-banda con detalles verticales (LH) y una sub-banda con detalles diagonales (HH) los cuales representan la información menos importante de la señal original [6-9], así mismo, esta descomposición se puede llevar a cabo en varios niveles [7-9].

La técnica de inserción “Quantization Index Modulation” propuesta en [10] se basa en cuantificar la señal portadora con cuantificadores que representan un símbolo de la marca de agua, en el caso de utilizar una marca de agua binaria se utilizan únicamente dos cuantificadores. La ecuación (1) muestra de manera general la cuantificación de una señal  $x$  por medio del cuantificador  $q(\cdot)$  el cual depende de las entradas  $m$  y  $\Delta$ , los cuales representan la marca de agua y el tamaño del paso de cuantificación respectivamente.

$$x' = q(x, m, \Delta) \quad (1)$$

La modulación dither [11-12] se refiere a modular la señal portadora con ruido pseudo-aleatorio generado de manera intencional, este ruido representa los vectores dither; se debe generar un vector dither  $d(m)$  por cada símbolo que compone la marca de agua. El primer vector dither utilizado definido como  $d[n, 0]$  es generado como un vector pseudo-aleatorio utilizando una llave o semilla, la cual es propuesta por [11-12], con una distribución uniforme de  $[-\Delta/2, \Delta/2]$  para  $n = 1, 2, 3, \dots, N$  donde  $N$  representa la longitud de los valores de la marca de agua a insertar. La generación de otro vector dither se realiza mediante (2) el cual se basa en  $d[n, 0]$ .

$$d[n, 1] = \begin{cases} d[n, 0] + \Delta/2, & d[n, 0] < 0 \\ d[n, 0] - \Delta/2, & d[n, 0] > 0 \end{cases} \quad n = 1, 2, 3, \dots, N \quad (2)$$

La inserción a través de QIM-DM se realiza de acuerdo a (3) para  $n = 1, 2, 3, \dots, N$ .

$$x'_n = q(x_n + d[n, m_n], \Delta) - d[n, m_n] \quad (3)$$

Para realizar la extracción de la marca de agua nuevamente se deben calcular los vectores dither con el mismo vector pseudo-aleatorio generado en la inserción, se realiza nuevamente la cuantización en (3) para cada vector dither y se calcula la distancia Euclidiana entre  $x'_n$  y  $x_n$ . Para el caso de una inserción binaria se realiza la extracción mediante (4).

$$\text{if } E(d[n, 0]) < E(d[n, 1]) \text{ } m_n = 0 \quad (4)$$

$$\text{else } m_n = 1$$

donde  $E(\cdot)$  representa la distancia Euclidiana entre  $x'_n$  y  $x_n$ .

La inserción de la marca de agua binaria del sistema propuesto se realiza en la sub-banda LL del segundo nivel de descomposición Wavelet por el método QIM-DM. Primero se realiza la extracción de una región en la cual se inserta la marca de agua seleccionada, esto se lleva a cabo para reducir el tiempo de procesamiento de la inserción y la extracción de la marca de agua. Se evaluó la robustez del sistema ante ataques de compresión JPEG y ante adición de ruido Gaussiano e Impulsivo. Por último se implementó el algoritmo propuesto utilizando computo en paralelo con CPU multi-núcleo y se obtuvieron tiempos de procesamiento del algoritmo de

inserción y extracción tanto de los algoritmos seriales como paralelos.

## II. ALGORITMO DE INSERCIÓN DE LA MARCA DE AGUA

La figura 1 muestra el diagrama a bloques del sistema propuesto para la inserción de la marca de agua. El primer paso consiste en leer la imagen portadora en formato RGB llamada  $I^{RGB}$  de tamaño  $M \times N$  donde  $M, N$  representan las filas y columnas de la imagen respectivamente. Se convierte la imagen  $I^{RGB}$  al espacio de color YCbCr obteniendo  $I^{YCbCr}$ , posteriormente se obtiene el canal de luminancia  $I^Y$  y se procede a extraer la región donde se insertará la marca de agua a partir de  $I^Y$ .

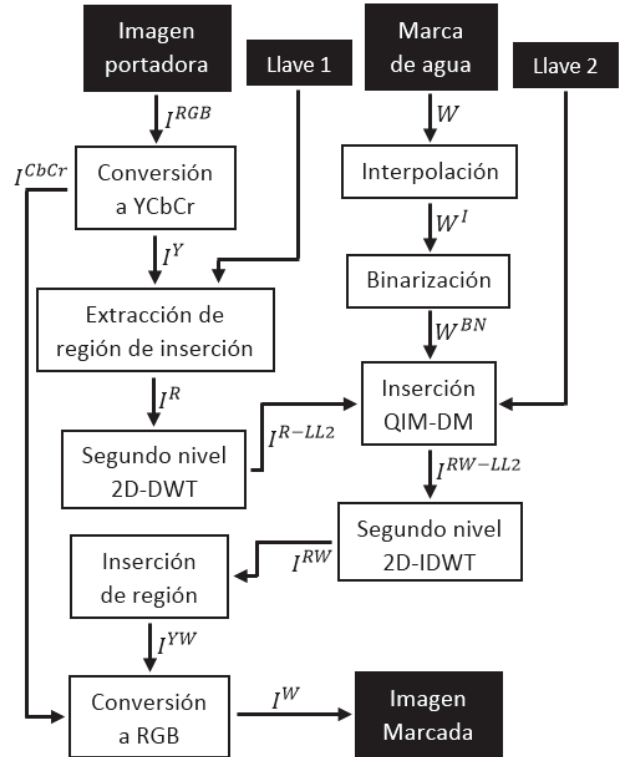


Figura 1. Diagrama a bloques del proceso de inserción

Para realizar el proceso de extracción de una región de inserción se calcula primero el número de píxeles en filas y columnas que deberá tener la marca de agua a ser insertada mediante (5) y (6).

$$\text{cant}X = \lfloor N/(\omega * \sigma) \rfloor \quad (5)$$

$$\text{cant}Y = \lfloor M/(\omega * \sigma) \rfloor \quad (6)$$

donde  $(\omega * \sigma)$  indica el tamaño de los bloques no traslapados extraídos de  $I^Y$  en donde en cada bloque se insertará un dato de la marca de agua,  $\sigma$  obtenido por (7) donde  $\epsilon$  representa el total de píxeles ocupados por las transformadas wavelet, en este caso al utilizar dos transformadas wavelet haar se obtiene un valor  $\epsilon = 16$  y  $\omega$  representa la cantidad de  $\sigma$  utilizadas para ocultar un dato de la marca de agua para  $\omega \geq 2$ .

$$\sigma = \sqrt{\epsilon} \quad (7)$$



Para obtener la región de inserción, llamada  $I^R$  de tamaño  $[cantY * \sigma, cantX * \sigma]$ , a partir de  $I^Y$  se extrae  $I^{Aux}$  de dimensiones  $[\omega * \sigma, \omega * \sigma]$  de acuerdo a (8).

$$I^{Aux} = I^Y(y * \omega * \sigma, x * \omega * \sigma) \quad (8)$$

Se extrae a partir de  $I^{Aux}$  un bloque de tamaño  $[\sigma, \sigma]$  y se asigna a la correspondiente posición de  $I^R$  de acuerdo a (9).

$$I^R(y * \sigma, x * \sigma) = I^{Aux}(\alpha, \beta) \quad (9)$$

donde  $\alpha$  y  $\beta$  se obtienen a partir de (10) y (11).

$$\alpha = \lfloor (\delta[cantX * y + x] - 1) / \omega \rfloor * \sigma \quad (10)$$

$$\beta = ((\delta[cantX * y + x] - 1) \% \omega) * \sigma \quad (11)$$

donde  $\delta$  representa un vector de tamaño  $[cantX * cantY]$  con valores pseudo-aleatorios generados por una llave o semilla.

Se realizan los pasos representados en (8), (9) (10) y (11) para  $x = 0, 1, 2, \dots, cantX - 1$  y para  $y = 0, 1, 2, \dots, cantY - 1$ . La figura 2 muestra el resultado del proceso de extracción de una región propuesta; en la cual 2(a) es el canal de luminancia  $I^Y$  de una imagen de tamaño 740x628 pixeles, 2(b) muestra la región de inserción para  $\omega = 2$  y 2(c) muestra la región de inserción para  $\omega = 4$ .

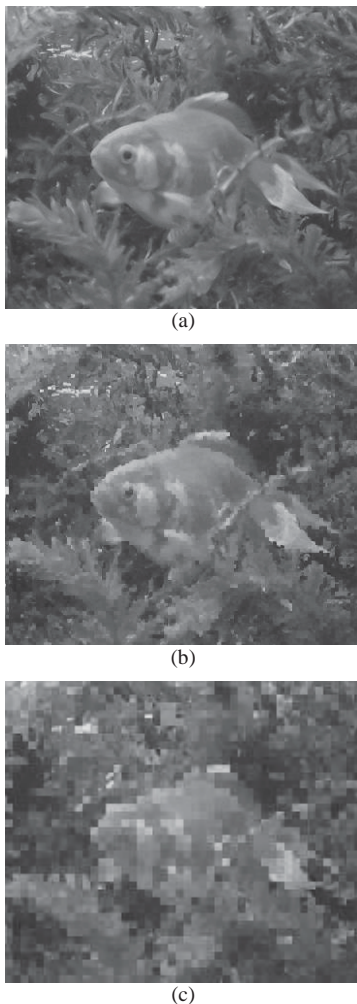


Figura 2. Extracción de region de inserción (a)  $I^Y$  740x628 pixeles (b)  $I^R$  para  $\omega = 2$  de 368x312 pixeles (c)  $I^R$  para  $\omega = 4$  de 184x156 pixeles.

Una vez obtenido  $I^R$  se aplica la 2D-DWT haar y a la sub-banda  $LL1$  se le aplicó nuevamente la 2D-DWT haar para obtener  $I^{R-LL2}$ .

Por otra parte se realiza un pre-procesamiento a la imagen que representa la marca de agua llamada  $W$  en el espacio de color RGB, se aplica una interpolación a  $W$  para que tenga las dimensiones de  $[cantY, cantX]$  obteniendo  $W^I$ , en este caso se utilizó la interpolación bicubica [13]. Posteriormente se binariza la imagen utilizando un umbral  $T$  para obtener  $W^{BN}$ , en el sistema propuesto se utilizó un valor de umbral  $T = 100$ .

Una vez obtenido  $I^{R-LL2}$  y  $W^{BN}$  se realiza la inserción de la marca  $W^{BN}$  dentro de  $I^{R-LL2}$  por medio del algoritmo QIM-DM utilizando (3) obteniendo  $I^{RW-LL2}$ . Se aplican los dos niveles de la 2D-IDWT haar para obtener  $I^{RW}$  y se aplica el proceso inverso de (8), (9) (10) y (11) para  $x = 0, 1, 2, \dots, cantX - 1$  y para  $y = 0, 1, 2, \dots, cantY - 1$  insertando la región marcada dentro de  $I^Y$  obteniendo  $I^{YW}$ . Por último tomando  $I^{YW}$  e  $I^{CbCr}$  se convierte la imagen del espacio de color YCbCr al espacio de color RGB para obtener  $I^W$  que representa la imagen marcada.

### III. ALGORITMO DE EXTRACCIÓN DE LA MARCA DE AGUA

La figura 3 muestra el diagrama a bloques del sistema propuesto para la extracción de la marca de agua. Inicialmente se lee la imagen marcada  $I^{RGB}$  de tamaño  $M \times N$  donde  $M, N$  representan las filas y columnas de la imagen respectivamente, la cual pudo haber sido sometida a ataques con la finalidad de remover la marca de agua. Se convierte del espacio de color RGB al espacio de color YCbCr y se extrae el canal de luminancia  $I^Y$ . Se aplica el proceso de extracción de la región donde se encuentra insertada la marca de agua llamada  $I^R$  de acuerdo a (8), (9) (10) y (11) para  $x = 0, 1, 2, \dots, cantX - 1$  y para  $y = 0, 1, 2, \dots, cantY - 1$ .

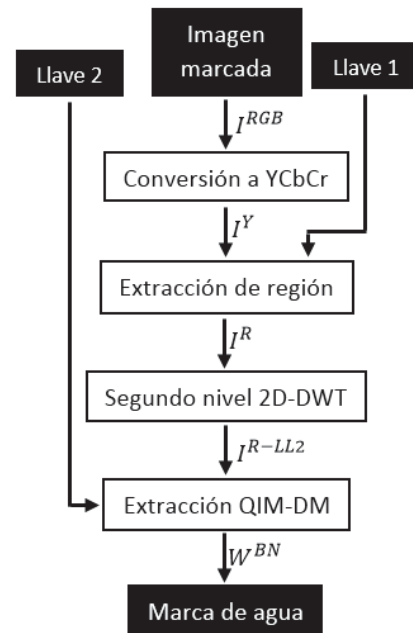


Figura 3. Diagrama a bloques del proceso de extracción

A la región obtenida  $I^R$  se aplica la 2D-DWT haar y a la sub-banda  $LL1$  se le aplicó nuevamente la 2D-DWT haar para obtener  $I^{R-LL2}$ . Utilizando el algoritmo QIM-DM inverso expuesto en (3) y (4) se extrae la marca de agua insertada  $W^{BN}$ .

#### IV. RESULTADOS

El sistema fue desarrollado en C++ utilizando Visual Studio 2010 Ultimate. Dentro de las pruebas realizadas se utilizaron imágenes portadoras de tamaño promedio de  $2592 \times 1944$  píxeles, las figuras 4 y 5 muestran un ejemplo de una imagen portadora y una de marca de agua de  $500 \times 802$  píxeles respectivamente.



Figura 4. Imagen portadora utilizada.



Figura 5. Marca de agua utilizada.

La figura 6 muestra las distintas marcas de agua generadas para  $\omega = 2$ ,  $\omega = 3$  y  $\omega = 4$ .



Figura 6. Marcas de agua utilizadas. (a)  $\omega = 2$ ;  $324 \times 243$  píxeles (b)  $\omega = 3$ ;  $216 \times 162$  píxeles (c)  $\omega = 4$ ;  $162 \times 121$  píxeles

Las primeras pruebas que se realizaron en el sistema fue calcular el PSNR (Relación Señal a Ruido Pico) que se obtiene por (12) y (13) comparando la imagen marcada contra la imagen original para evaluar la imperceptibilidad de la marca de agua binaria.

$$MSE = \frac{1}{X \cdot Y \cdot Z} \sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} \sum_{k=0}^{Z-1} [I(i, j, k) - I'(i, j, k)]^2 \quad (12)$$

$$PSNR = 10 \log_{10}(255^2 / MSE) \quad (13)$$

Se calculó la Similitud Estructural (SSIM) obtenido por (14).

$$SSIM(x, y) = [L(x, y)^\alpha \cdot C(x, y)^\beta \cdot S(x, y)^\gamma] \quad (14)$$

Dónde los parámetros  $L$ ,  $C$  y  $S$  se definen mediante (15), (16) y (17).

$$I(x, y) = (2\mu_x\mu_y + C_1) / (\mu_x^2 + \mu_y^2 + C_1) \quad (15)$$

$$C(x, y) = (2\sigma_x\sigma_y + C_2) / (\sigma_x^2 + \sigma_y^2 + C_2) \quad (16)$$

$$S(x, y) = (\sigma_{xy} + C_3) / (\sigma_x + \sigma_y + C_3) \quad (17)$$

donde  $\mu$  y  $\sigma$  son la media y la varianza de las imágenes  $X$  y  $Y$  respectivamente,  $\sigma_{xy}$  representa la covarianza de  $X$  y  $Y$ ;  $C_1$ ,  $C_2$ ,  $C_3$  son constantes muy pequeñas. Tomando  $\alpha=\beta=\gamma=1$  y  $C_3=C_2/2$  la ecuación se reescribe como ec. (18) [14].

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_x\sigma_y + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (18)$$

La figura 7 muestra el PSNR obtenido para diferentes valores de  $\Delta$  (Step) utilizando  $\omega = 2$ ,  $\omega = 3$  y  $\omega = 4$ . Como se puede observar en la figura 6 se obtienen buenos resultados en cuanto a imperceptibilidad de la marca de agua ya que el PSNR se encuentra por arriba de 35dB.

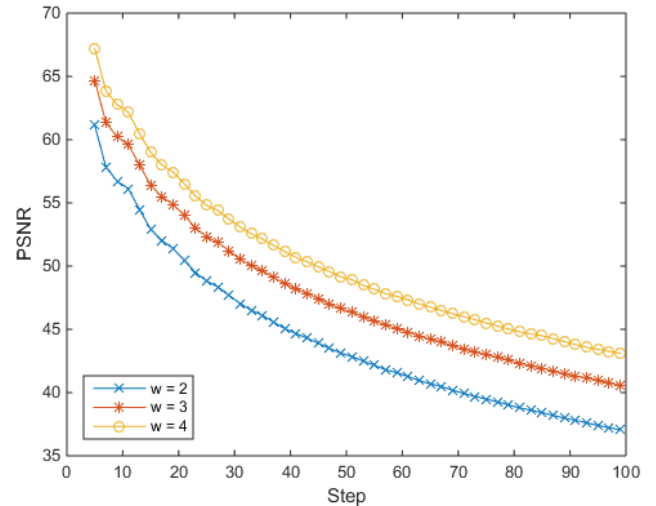


Figura 7. PSNR de imagen marcada contra imagen original.

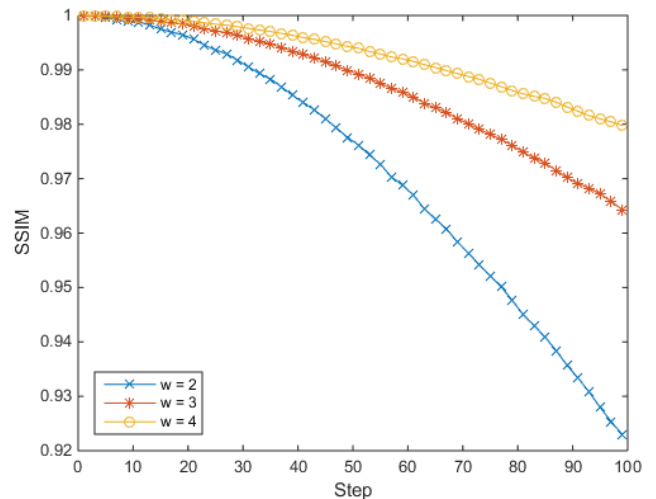


Figura 8. SSIM de imagen marcada contra imagen original.

La figura 8 muestra el SSIM obtenido para diferentes valores de  $\Delta$  (Step) utilizando  $\omega = 2$ ,  $\omega = 3$  y  $\omega = 4$ . En este caso se observa como la similitud entre la imagen original y la imagen marcada decae con un valor de  $\omega$  menor ( $\omega = 2$ ). De acuerdo a las figuras 6 y 7 se toman como posibles  $\Delta$  los valores entre 20 y 80 debido a la buena calidad de la imagen marcada.

Para la fase de extracción de la marca de agua, el sistema fue sometido a ataques de compresión JPEG e inserción de ruido impulsivo y gaussiano. Para evaluar la calidad de la marca de agua extraída se calculó el coeficiente de correlación (NC) por medio de (19), (20) y (21) comparando la marca de agua insertada contra la marca de agua extraída ( $W^{BN}$  para ambos casos).

La figura 9 muestra el NC obtenido después del ataque de compresión JPEG utilizando un valor  $\omega = 3$  para distintos factores de compresión. Como se puede observar, el sistema propuesto tiene robustez ante el ataque de compresión para valores  $\Delta \geq 40$  ya que soporta factores de compresión de hasta 30 obteniendo buena calidad en la marca de agua extraída.

$$NC = \frac{\sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} [W(i,j) - \bar{W}][W'(i,j) - \bar{W}']}{\sqrt{(\sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} [W(i,j) - \bar{W}]^2)(\sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} [W'(i,j) - \bar{W}']^2)}} \quad (19)$$

$$\bar{W} = \frac{\sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} W(i,j)}{X*Y} \quad (20)$$

$$\bar{W}' = \frac{\sum_{i=0}^{X-1} \sum_{j=0}^{Y-1} W'(i,j)}{X*Y} \quad (21)$$

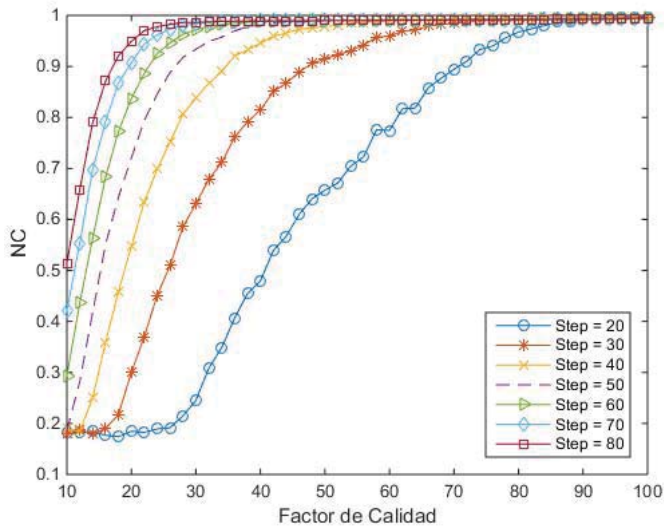


Figura 9. NC obtenido para ataque de compresión JPEG, utilizando  $\omega = 3$ .

La figura 10 muestra el NC obtenido después del ataque de adición de ruido impulsivo utilizando un valor  $\omega = 3$ . Como se puede observar la robustez para este tipo de ruido es muy parecida para cada valor de  $\Delta$  debido a que el ruido impulsivo únicamente ataca ciertas partes de la imagen, por lo tanto la robustez a este ataque depende del valor de  $\omega$  utilizado.

Para un valor de  $\omega = 3$ , el sistema propuesto es robusto hasta una densidad de 0.01 de ruido, lo cual se podría considerar bueno ya que con esta densidad la imagen se altera de manera significativa. Véase figura 11.

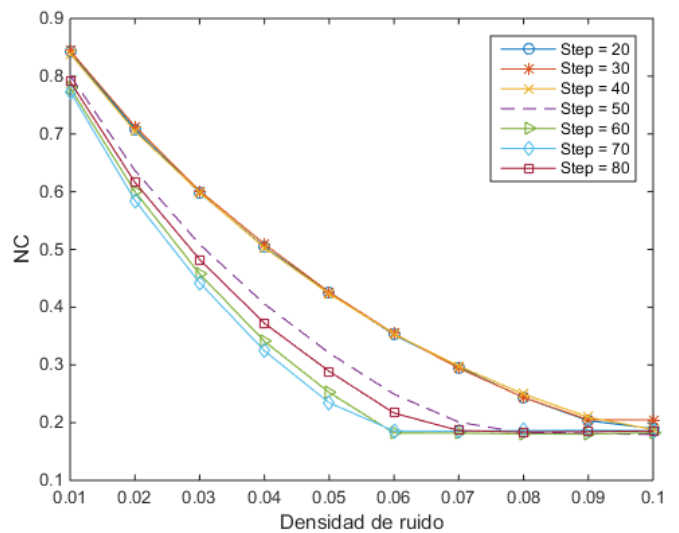


Figura 10. NC obtenido para ataque de ruido Impulsivo, utilizando  $\omega = 3$ .



Figura 11. Ataque de ruido impulsivo con densidad de 0.01.

La figura 12 muestra el NC obtenido después del ataque de adición de ruido Gaussiano utilizando un valor  $\omega = 3$ . Se puede observar que la robustez ante el ruido Gaussiano depende del valor de  $\Delta$ , para  $\Delta = 80$  soporta adición de ruido Gaussiano con varianza de 0.0015 lo cual es un buen resultado ya que utilizando este valor se altera de manera significativa la calidad de la imagen. Véase figura 13.

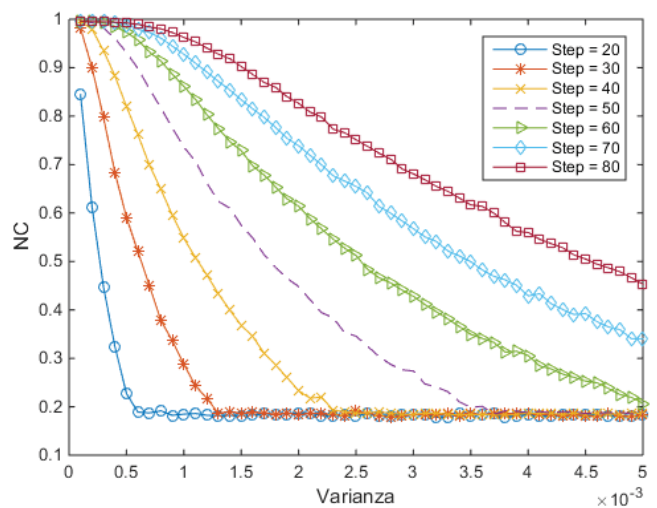


Figura 12. NC obtenido para ataque de ruido Gaussiano, utilizando  $\omega = 3$ .

Se obtuvieron tiempos de procesamiento para el algoritmo de inserción y extracción de la marca de agua utilizando dos equipos, se ejecutó el algoritmo de inserción y extracción 100 veces y se calculó el promedio del tiempo de procesamiento en segundos. La tabla 1 muestra las características de los equipos utilizados durante las pruebas. La tabla 2 y la tabla 3 muestran los tiempos de procesamiento ejecutando el algoritmo de manera secuencial y ejecutando el algoritmo de manera paralela utilizando CPU multi-núcleo para los equipos 1 y 2 respectivamente. Para la ejecución del algoritmo de manera paralela se utilizó la librería OpenMP en C++.



Figura 13. Ataque de ruido Gaussiano con varianza de 0.0015.

Tabla 1. Características de los equipos utilizados.

Elemento	Equipo 1	Equipo 2
CPU	AMD FX-6300 3.5GHz (6 núcleos)	Intel Celeron 2.16GHz (2 núcleos)
RAM	8 GB DDR3	4 GB DDR3
Disco Duro	1TB SATA II	500GB SATA II
Sistema Operativo	Windows 7 Ultimate	Windows 8.1 Pro

Tabla 2. Tiempos de procesamiento en segundos para el equipo 1.

$\omega$	Serial (6 núcleos)		Serial (1 núcleo)		Paralelo (6 núcleos)	
	Inserción	Extracción	Inserción.	Extracción	Inserción	Extracción
2	2.787	1.262	2.845	1.397	1.594	0.641
3	1.345	0.628	1.557	0.647	1.017	0.358
4	1.094	0.405	1.156	0.418	0.595	0.257

Tabla 3. Tiempos de procesamiento en segundos para el equipo 2.

$\omega$	Serial (2 núcleos)		Serial (1 núcleo)		Paralelo (2 núcleos)	
	Inserción	Extracción	Inserción.	Extracción	Inserción	Extracción
2	4.921	2.255	4.953	2.357	3.927	2.152
3	2.815	1.044	3.711	1.136	2.216	0.838
4	2.126	0.811	2.252	0.953	2.055	0.613

Como se puede observar en las tablas 2 y 3 el sistema fue ejecutado de manera serial sin utilizar la librería OpenMP (columnas 2 y 3) para observar los tiempos de procesamiento en los cuales el algoritmo se ejecutaría al no aprovechar el multi-núcleo de los procesadores actuales. En las columnas 4 y 5 se observan los tiempos de procesamiento al ejecutar el algoritmo utilizando OpenMP y activando únicamente 1 núcleo del procesador para que el sistema se ejecute de manera serial. Como se puede observar la reducción de tiempo de procesamiento es mínima comparando las ejecuciones serial utilizando todos los núcleos y utilizando un solo núcleo del procesador. Las columnas 6 y 7 muestran los tiempos de procesamiento ejecutando el algoritmo con cómputo en paralelo utilizando todos los núcleos que posee el procesador.

## V. CONCLUSIONES

En este trabajo se propone y desarrolla un sistema computacional para la protección de derechos de autor utilizando DWT para transformar la imagen portadora al dominio de la frecuencia, utilizando la técnica QIM-DM se inserta la marca de agua binaria en la sub-banda LL del segundo nivel de descomposición Wavelet. Es posible utilizar otros niveles de descomposición pero se podría ver afectado la capacidad de inserción y de la misma manera la robustez del sistema. Por otra parte se utilizó un algoritmo para extraer la región en la cual se insertaría la marca de agua, dicha región depende de un valor  $\omega$  el cual indica el tamaño de la misma, utilizando vectores pseudo-aleatorios se extrae la región. El sistema utiliza dos llaves las cuales dan mayor seguridad al algoritmo para evitar técnicas de extracción de la marca de agua.

El tamaño de la marca de agua a ser utilizada y los tiempos de procesamiento del sistema dependen de  $\omega$ , el cual puede ser modificado de acuerdo a los requerimientos del usuario.

De acuerdo a la robustez del sistema ante ataques de compresión y adición de ruido, el valor  $\Delta$  recomendable es 80 aunque este podría aumentar si se requiere aumentar la robustez del sistema siempre y cuando la calidad de la imagen marcada no sea menor a 35dB.

Utilizando cómputo en paralelo el sistema podría ser utilizado para la inserción y extracción de una marca de agua en video digital utilizando  $\omega = 2$  debido a que el sistema es mucho más rápido en comparación a ejecutarlo de manera secuencial.

## AGRADECIMIENTOS

Agradecemos al Instituto Politécnico Nacional, al CONACYT, a la COFAA y a la Beca de Estimulo Institucional de Formación de Investigadores (BEIFI) del IPN por el apoyo otorgado para el desarrollo de este trabajo.

## REFERENCIAS

- [1] Manuel García Horta. Autenticación de documentos digitales usando la técnica de marca de agua. Escuela Superior de Ingeniería Mecánica y Eléctrica Unidad Culhuacan, SEPI ESIME Culhuacan, IPN 2010.
- [2] Izlian Yolanda Orea Flores. Marcas de agua robustas en imágenes digitales con formato BMP. Escuela Superior de Ingeniería Mecánica y Eléctrica Unidades Zacatenco, SEPI ESIME Zacatenco, IPN 2005.
- [3] Jialinh Han, Xiaohui Zhao "A Novel Watermarking Method Based on Watermark Compression Preprocessing" 2013 Fourth International Conference on Emerging Intelligent Data and Web Technologies (EIDWT), Xi'an, 2013, pp. 676-679.
- [4] Zhi-Feng Gu, Guan-Feng Li, Zong-Xiao Yang "Study on digital image watermark algorithm based on chaos mapping and DWT," 2012 International Conference on Advanced Mechatronic Systems (ICAMechS). Tokyo, 2012, pp. 160 – 164.
- [5] Xueyi Ye, Xueting Chen, Meng Deng, Shuyun Hui, Yunlu Wang, "A Multiple-Level DCT Based Robust DWT-SVD Watermark Method," 10th International Conference on Computational Intelligence and Security, 2014, China, pp. 479 – 483.
- [6] Irshad Ahmad Ansari, Millie Pant, Ferrante Neri, "Analysis of Gray Scale Watermark in RGB Host using SVD and PSO," IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing (CIMSIVP), 2014, Orlando, FL, pp. 1-7.
- [7] Sanjay R. Zadokar, Vishal B. Raskar, Sachin V. Shinde, "A Digital Watermarking for Anaglyph 3D images," International Conference on

- Advances in Computing, Communications and Informatics (ICACCI). 2013, pp. 483 - 488.
- [8] D. Usha, Y. Rakesh, Generation of Digital Watermarked Anaglyph 3D Image Using DWT. SSRG International Journal of Electronics and Communication Engineering (SSRG-IJECE) – volume1 issue7. 2014, pp. 33 - 37.
- [9] Jumana Waleed, H.Dong Jun, Saad Hameed, Hiyam Hatem, Raed Majeed, “Integral Algorithm to Embed Imperceptible Watermark into Anaglyph 3D Video,” International Journal of Advancements in Computing Technology (IJACT), Volume5, Number13. 2013, pp. 163 – 173.
- [10] B. Chen, G. W. Wornell, “Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding,” IEEE Transaction on Information Theory, vol. 47, No. 4, 2001, pp. 1423-1443.
- [11] Qiao Li, Ingemar J. Cox, “Rational Dither Modulation Watermarking Using a Perceptual Model,” IEEE 7th Workshop on Multimedia Signal Processing, Shanghai , 2005, pp. 1-4
- [12] Brian Chen, Gregory W. Wornell, “Quantization Index Modulation Methods for Digital Watermarking and Information Embedding of Multimedia,” Journal of VLSI Signal Processing. Kluwer Academic Publishers. Manufactured in The Netherlands, 2001, pp. 7-33.
- [13] Nuno-Maganda, M.A., Arias-Estrada, M.O., “Real-time FPGA-based architecture for bicubic interpolation: an application for digital image scaling,” ReConFig. International Conference on Reconfigurable Computing and FPGAs, Puebla, 2005.
- [14] Su, K., Xiao, C., & Ye, S. “Video Quality Assessment Based on Edge Structural Similarity,” Congress on Image and Signal Processing IEEE computer society, 2008

# Inserción de una Marca de Agua Visible Adaptativa basada en la Luminancia y Textura de la Imagen Digital Huésped a Color

*Kevin Rangel Espinoza<sup>1</sup>, Clara Cruz Ramos, Rogelio Reyes Reyes, Mariko Nakano Miyatake.*  
 Maestría en Ingeniería en Seguridad y Tecnologías de la Información  
 Instituto Politécnico Nacional, ESIME Unidad Culhuacán  
 Av. Santa Ana No. 1000, Col. San Francisco Culhuacán, CP. 04430, México, D.F.  
 Tel. y Fax (55) 56562058 email: kevinrespinoza@gmail.com<sup>1</sup>

**Resumen**— La mayoría de los algoritmos de marca de agua visibles que se encuentran dentro de la literatura, no consideran las características de textura y luminancia de la imagen huésped para realizar la inserción de la marca de agua. Para lograr un equilibrio óptimo entre la calidad visual de la imagen huésped y la fuerza de inserción de la marca de agua, un esquema de inserción de marca visible adaptativa la cual considera características del Sistema Visual Humano (SVH) es propuesta en este trabajo; donde el factor de escalamiento y de inserción de la marca de agua se calculan mediante la combinación de las características de luminancia y textura de la imagen huésped. Los resultados experimentales demuestran que la imagen marcada tiene una buena calidad perceptible y alta robustez frente a los ataques intencionales para remover la marca de agua.

**Palabras Clave**—*Marca de Agua Visible, Sistema Visual Humano, Factor de Escalamiento, Factor de inserción, Marca de Agua Adaptativa.*

## I. INTRODUCCIÓN

En los últimos años, las marcas de agua digitales han sido propuestas como una solución para la protección de derechos de autor de contenido multimedia, por lo que ha generado un amplio interés dentro de los investigadores de diferentes campos. Las marcas de agua se han convertido en un tema de interés dentro de las Tecnologías de la Información (TI) [1]. En general, las marcas de agua digitales se pueden clasificar en dos grandes grupos visibles e invisibles. Las marcas de agua invisibles insertan información dentro del contenido digital, de tal manera que ésta no pueda ser visualizada a simple vista por lo que es necesario el uso de algoritmos para la extracción de la información; mientras que en las marcas de agua visibles la información que es insertada es intencionalmente perceptible para el ojo humano.

Dentro de la literatura de las marcas de agua, muy pocos investigadores se centran en las marcas de agua visibles, esto se debe a que, en algunos casos, una marca de agua visible reducirá el valor comercial del contenido marcado. De hecho, las marcas de agua visibles pueden ser aplicadas en diferentes campos tales como bibliotecas digitales, comercio electrónico y televisión digital [2]. Una marca de agua visible eficiente debe de contener las siguientes características [3]:

- La marca debería poder apreciarse tanto en imágenes a escalas de grises como en imágenes a color, esto indica

que la información de la marca de agua visible deberá ser insertada dentro de la luminancia.

- La marca de agua debe de insertarse en un área grande o importante de la imagen con el fin de poder evitar su eliminación ante recorte.
- La marca de agua deberá ser apreciable a simple vista, pero sin obstruir de manera considerable los detalles de la imagen huésped.
- La marca de agua debe ser difícil de eliminar; la remoción de la marca de agua debe ser un trabajo laborioso y más costoso que la compra de la imagen a través del propietario.

Hasta ahora existen muy pocos esquemas de marca de agua visible reportados dentro de la literatura. La organización de librería digital IBM utiliza técnicas de marcado de agua basadas en el dominio espacial para poder proteger manuscritos digitales del Vaticano. Kankanhalli y Rajmohan [4] propusieron una técnica de marca de agua visible en el dominio de la Transformada Discreta del Coseno (DCT), en el que la fuerza de inserción de la marca de agua varía de acuerdo al contenido de la imagen huésped. En esta técnica la sensibilidad de la distorsión de los bloques es calculada por el análisis de la textura, borde y luminancia de cada uno de ellos. Mohanty y Ramakrishna [5] mejoraron la técnica de Kankanhalli y Rajmohan al fortalecer la adaptabilidad de la marca de agua al considerar las características de la marca de agua al momento de su inserción. Yang [6] presentó un esquema de marca de agua visible removible que puede ser aplicable en diferentes campos, este esquema consiste en la inserción de una marca de agua visible que puede ser removida por el usuario autorizado mediante el uso de una llave para así poder obtener la imagen original. El algoritmo de marcado de agua inserta la marca de agua en una región especificada por el usuario dentro de la imagen huésped mediante el ajuste adaptativo del valor de los píxeles de la imagen huésped y la marca de agua considerando la sensibilidad del ojo humano al contraste. Tsai [7] presentó un algoritmo de marcado de agua visible considerando la sensibilidad al contraste de la imagen huésped. Este algoritmo es capaz de obtener las ubicaciones óptimas dentro de la imagen huésped para la inserción de la marca de agua, así como la fuerza de inserción considerando las características de textura y luminancia de la imagen huésped y la marca de agua en la Transformada Discreta de Wavelet (DWT). Para obtener un

equilibrio óptimo entre la calidad visual de la imagen marcada y la visibilidad de la marca de agua, es necesario tomar en cuenta las características de luminancia y textura de la imagen huésped, basado en las características del SVH [8]:

- El ojo humano posee diferentes grados de sensibilidad a diferentes luminancias, pero generalmente este es más sensible a la luminancia media.
- El ojo humano es más sensible al ruido de la imagen en zonas lisas que en zonas altamente texturizadas.

## II. SISTEMA PROPUESTO

### A. Generación de la Marca de Agua

Para la inserción de la marca de agua visible es necesario calcular los valores de escalamiento e inserción los cuales son obtenidos de acuerdo a las características del SVH [9]. La Ec. (1) es utilizada para la inserción de la marca de agua visible.

$$I_m(i, j) = \alpha_m I_m^h(i, j) + \beta_m I_m^w(i, j) \quad (1)$$

$i = 1, 2 \dots, 8, j = 1, 2 \dots, 8, m = 1, 2 \dots, M.$

Donde  $I_m(i, j)$ ,  $I_m^h(i, j)$  e  $I_m^w(i, j)$  denotan el  $i, j$ -ésimo coeficiente en el dominio de la DCT del  $m$ -ésimo bloque de  $8 \times 8$  de la imagen marcada  $I$ , la imagen huésped  $I^h$  y la marca de agua  $I^w$  respectivamente.  $\alpha_m$  y  $\beta_m$  son los factores de escalamiento e inserción para cada bloque  $m$  de la imagen huésped  $I_m^h$  y la marca de agua  $I_m^w$  respectivamente.  $M$  es el número total de bloques de tamaño  $8 \times 8$  en las imágenes. El algoritmo de inserción descrito anteriormente se aplica por separado en cada uno de los canales del espacio de color RGB.



Figura 1.- Canales Rojo, Verde y Azul que componen una imagen digital.

### B. Calculo de los factores de escalamiento $\alpha_m$ e inserción $\beta_m$

Los valores de escalamiento e inserción  $\alpha_m$  y  $\beta_m$  de la Ec. (1) son utilizados para determinar la fuerza de la imagen huésped y de la marca de agua, esto determina la visibilidad de la marca de agua y la robustez de la imagen marcada. El algoritmo utilizado toma las características del SVH considerando las texturas y luminancia de la imagen huésped  $I^h$ . Hay dos aspectos que hay que tomar en cuenta para la generación del factor de escalamiento  $\alpha_m$ :

- Primero el SVH es muy sensible a los cambios de luminancia, es decir para no alterar de manera significativa la imagen huésped se le debe de asignar un valor al factor de escalamiento  $\alpha_m$  tomando en consideración la luminancia promedio, maximizando este valor en estas componentes y disminuyéndolo en las componentes de luminancia más oscuras y más brillantes, acercándose a una forma gaussiana; esto con la finalidad de preservar la luminancia general de la imagen huésped original sin degradar su calidad.
- En segundo lugar el SVH es menos sensible a las regiones altamente texturizadas, por lo que se debe de utilizar un valor más alto del factor de escalamiento  $\alpha_m$  en regiones texturizadas, lo que preserva las características de bordes y regiones planas de la imagen original, lo que origina que se preserve la calidad de la imagen huésped [10].

Los pasos para determinar el factor de escalamiento  $\alpha_m$  y el factor de inserción  $\beta_m$  son los siguientes:

Paso 1: Convertir la Imagen huésped  $I^h$  a formato RGB y trabajar sobre los canales Rojo  $I^{hr}$ , Verde  $I^{hg}$  y Azul  $I^{hb}$ .

Paso 2: Convertir la Marca de Agua  $I^w$  a formato RGB y trabajar sobre los canales Rojo  $I^{wr}$ , Verde  $I^{wg}$  y Azul  $I^{wb}$ .

Paso 3: Dividir los Canales Rojo, Verde y Azul de la imagen huésped  $I^h$  y la marca de agua  $I^w$  en bloques de  $8 \times 8$  pixeles y a cada bloque aplicarle la transformada DCT.

Paso 4: Debido a que la mayor parte de la luminancia de cada bloque es concentrada en los componentes de menor frecuencia, especialmente en los coeficientes DC, se calculara el factor de escalamiento  $\alpha_m$  de acuerdo a los coeficientes DC de cada canal de la imagen  $I^h$  y de la marca de agua  $I^w$ . La distribución de los coeficientes DC para cada canal de la imagen  $I^h$  y de la marca de agua  $I^w$  son expresados a continuación.

$$I_m^h(0,0): N(\mu_1, \sigma_1^2), \quad m = 1, 2, 3 \dots M \quad (2)$$

$$I_m^w(0,0): N(\mu_2, \sigma_2^2), \quad m = 1, 2, 3 \dots M \quad (3)$$

Donde  $\mu_1, \sigma_1^2$  y  $\mu_2, \sigma_2^2$  son la media y la varianza de los coeficientes DC para los canales RGB de la imagen  $I^h$  y la marca de agua  $I^w$  respectivamente. La variable  $V_m$  se calcula para cada canal de color RGB mediante la siguiente ecuación:  $V_m = I_m^h(0,0) + I_m^w(0,0)$   $m = 1, 2, 3 \dots M$ . Para calcular el factor de escalamiento se usa la siguiente ecuación:

$$\alpha_m = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{[V_m - (\mu_1 + \mu_2)]^2}{2(\sigma_1^2 + \sigma_2^2)}\right\} \quad (4)$$

Los valores de la media  $\mu_1$  y la varianza  $\sigma_1^2$  de los coeficientes DC para cada bloque de la imagen huésped están definidos como:

$$\mu_1 = \frac{1}{M} \sum_{m=1}^M I_m^h(0,0) \quad (5)$$

$$\sigma_1^2 = \frac{1}{M} \sum_{m=1}^M I_m^h(0,0) - \mu_1 \quad (6)$$

De la misma forma se calcula la media  $\mu_2$  y la varianza  $\sigma_2^2$  para cada canal de la marca de agua  $I^w$ . El procedimiento utilizado en la ec. (5) y la ec. (6), refleja una de las principales características del SVH debido a que se toman solo los coeficientes DC de cada bloque de los canales RGB, los cuales se encargan de transmitir la luminancia.

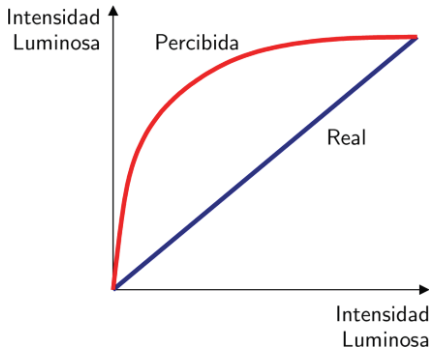


Figura 2.- Nivel de Luminancia percibida por el ojo humano.

Paso 5: Tomando en cuenta la segunda característica del SVH y con el fin de obtener un mejor factor de escalamiento, se tomarán en cuenta los coeficientes AC para los canales RGB debido a que estos coeficientes almacenan la energía de las texturas dentro de la imagen. Esto se debe a que la marca de agua  $I^w$  deberá de tener un mayor factor de inserción  $\beta_m$  en regiones altamente texturizadas debido a que el SVH es menos sensible a los cambios en estas regiones.

Del mismo modo el valor de inserción  $\beta_m$  es utilizado en las regiones de alta textura para asegurar la resistencia contra ataques de remoción ilegal de la marca de agua  $I^w$ . Asumiendo que el factor de escalamiento  $\alpha_m$  es directamente proporcional a la varianza  $x_m$  donde  $x_m = x_m^h + x_m^w$ . Los parámetros  $x_m^h$  y  $x_m^w$  son las varianzas de los coeficientes AC insignificantes para cada uno de los canales RGB de la imagen  $I^h$  y de la marca de agua  $I^w$  respectivamente. Un coeficiente AC es considerado insignificante si su valor de cuantización es igual a 0. Por lo tanto la ec. (4) puede describirse como:

$$\alpha_m = \frac{1}{\sqrt{2\pi(\sigma_1^2 + \sigma_2^2)}} \exp\left\{-\frac{|V_m - (\mu_1 + \mu_2)|}{2(\sigma_1^2 + \sigma_2^2)}\right\} + x_m^s \quad (7)$$

Donde  $x_m^s$  es la versión normalizada de  $x_m$  calculada de acuerdo a la ec. (8), en donde  $x_m^l$  es el logaritmo natural de  $x_m$ ,  $x_{min}^l$  y  $x_{max}^l$  son respectivamente los valores mínimos y máximos de todos los valores de  $x_m^l$ .

$$x_m^s = \frac{0.8 x (x_m^l - x_{min}^l)}{(x_{max}^l - x_{min}^l)} + 0.1 \quad (8)$$

$x_m^s$  es calculada en la ec. (9) dentro de un rango mínimo, esto sirve para estabilizar el valor de escalamiento  $\alpha_m$  sin alterar la calidad visual de la imagen marcada  $I$ . Los valores de la media  $n_m$  y la varianza  $x_m^h$  de los coeficientes AC para cada canal de la imagen huésped están definidos como:

$$x_m^h = \frac{1}{s_m^r} \sum_{(i,j) \in s_m^r} [I_m^h(i,j) - n_m]^2 \quad (9)$$

Donde  $s_m^r$  es la matriz de los coeficientes AC insignificantes y  $n_m$  es el valor de la media de estos. De la misma forma se obtiene la varianza  $x_m^w$  para la marca de agua  $I_m^w$ .

Paso 6: Calcular  $\beta_m = 1 - \alpha_m$  para cada uno de los canales; los factores de escalamiento  $\alpha_m$  e inserción  $\beta_m$  se interpolan de acuerdo a los siguientes rangos [0.78, 0.87] y [0.15, 0.20] respectivamente.

### C. Inserción de la marca de Agua Visible

Tomando en cuenta los canales RGB de la marca de agua  $I^w$ , los canales de la imagen huésped  $I^h$  y los valores de escalamiento  $\alpha_m$  e inserción  $\beta_m$ , los pasos a seguir son los siguientes:

Paso 1: Dividir los canales RGB de la imagen huésped  $I^h$  y la marca de agua  $I^w$  en bloques de 8x8 pixeles y a cada bloque aplicarle la transformada DCT.

Paso 2: Para cada bloque de los canales  $I^{hr}$ ,  $I^{hg}$ ,  $I^{hb}$  y la marca de agua  $I^{wr}$ ,  $I^{wg}$ ,  $I^{wb}$  generar los bloques de la imagen marcada  $I_m^r$ ,  $I_m^g$ ,  $I_m^b$  utilizando la ec. (1).

Paso 3: Realizar la transformada inversa DCT a cada bloque de los canales RGB  $I_m$  para producir la imagen marcada  $I$ .

## III. RESULTADOS EXPERIMENTALES

Para evaluar la eficiencia del sistema propuesto se sometió a varias pruebas con diferentes imágenes utilizadas comúnmente en la literatura de procesamiento de imágenes como Lenna, F-16, Baboon, etc. Las imágenes utilizadas son de tamaño 512x512 pixeles como imagen huésped.



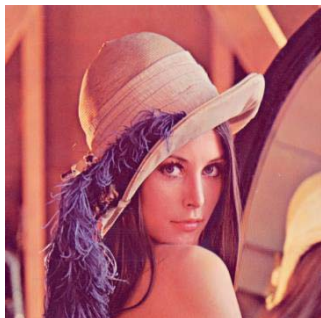


Figura 3.- Marca de Agua Visible

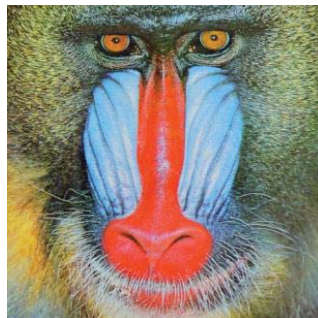
De igual forma se utilizó una marca de agua a color del mismo tamaño que la imagen huésped.

La marca de agua visible insertada en las imágenes digitales se muestra en la figura 5. Algunas de las imágenes utilizadas para la inserción de la marca de agua se muestran en la figura 4.

Como se alcanza a visualizar, las imágenes marcadas tienen altos valores de PSNR debido a que la fuerza de inserción de la marca de agua depende de las características de la imagen huésped. Los resultados sobre la transparencia de la marca de agua se presentan en la Tabla 1, el esquema propuesto tiene valores PSNR constantes para diferentes tipos de imágenes con valor promedio de 24.0 dB lo cual indica que el esquema propuesto puede lograr una buena calidad visual para diferentes tipos de imágenes huésped.



(a) Lenna

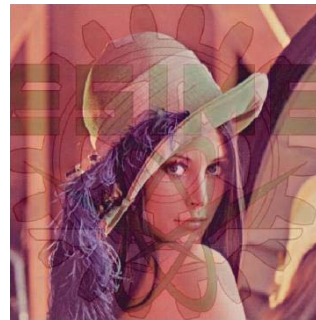


(b) Baboon

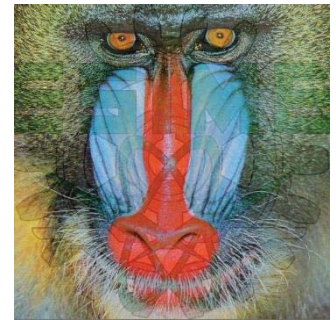


(c) F-16

Figura 4.- Imágenes Prueba.



(a) Lenna



(b) Baboon



(c) F-16

Figura 5- Imágenes Marcadas

Tabla 1. Transparencia de la Marca de Agua Visible

Imágenes	PSNR
Lenna	25.75
Baoon	26.23
F-16	24,12

En general la calidad visual de las imágenes marcadas con el esquema propuesto tiene mayor capacidad de control en la visibilidad de la marca de agua utilizando las características del SVH, a diferencia de hacer la inserción directamente con un factor de escalamiento que no toma en cuenta la luminocidad de la textura, como se muestra en la figura 6.

#### IV. CONCLUSIONES

El sistema de marca de agua visible adaptativo propuesto fue desarrollado explotando al máximo las características del sistema visual humano. Este esquema propuesto realiza la inserción de la marca de agua en el dominio de la frecuencia utilizando la transformada discreta del coseno. La fuerza de inserción de la marca de agua depende de las características de la imagen huésped y de la marca de agua, lo que ofrece como ventaja que la fuerza de inserción de la marca de agua sea menor en bloques planos y bordes, pero mayor en bloques altamente texturizados. Así también, la fuerza de inserción de la marca de agua es menor en bloques con luminancia promedio para evitar obstruir los rasgos característicos de la imagen huésped. Esta propuesta de inserción de la marca de agua adaptativa basada en el análisis de luminosidad y textura garantiza la preservación de la calidad original de la imagen huésped al mismo tiempo que se hace visible la marca de agua del propietario. Los resultados experimentales demostraron la

eficiencia del algoritmo al obtener valores promedio de PSNR de 25 dB, manteniendo una buena visibilidad de la marca de agua combinando luminancia y texturas, lo que conlleva que la marca de agua insertada sea robusta ante intentos de remoción ilegal, pero a la vez que no sea tan intrusiva con el contenido de la imagen huésped.

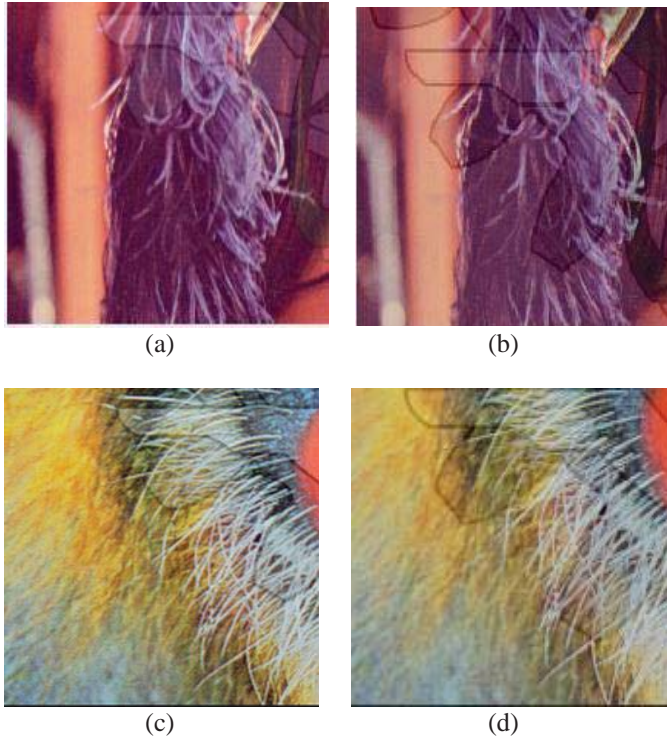


Figura 6.- (a),(c) Marca de Agua Visible Aditiva, (b), (d) Marca de Agua Visible Adaptativa Propuesta.

## V. AGRADECIMIENTOS

Al Instituto Politécnico Nacional, al Consejo Nacional de Ciencia y Tecnología y la Comisión de Operación y Fomento de Actividades Académicas y al BEIFI, por su apoyo y patrocinio para la realización de este trabajo.

## VI. REFERENCIAS

- [1] Braudaway, G.W. K.A. Magerlein and F.C. Mintzer, 1996. Protecting publicly available images with a visible image watermark. Proceedings of the SPIE Conference on Optical Security and Counterfeit Deterrence Technique, Feb. 1-2, San Jose, CA, USA., pp: 126-133.
- [2] Cox, I.J., G. Doerr and T. Furon, 2006. Watermarking is not cryptography. Proceedings of the 5th International Workshop on Digital Watermarking, Nov. 8-10, Jeju Island, Korea, pp: 1-15.
- [3]Huang, B. B., & Tang, S. X. (2006). A contrast-sensitive visible watermarking scheme. *IEEE MultiMedia*, (2), 60-66.
- [4]Kankanhalli, M.S. and R.K.R. Rajmohan, 1999. Adaptive visible watermarking of images. Proceedings of the IEEE International Conference on Multimedia Computing and Systems, July 7-11, Florence, Italy, pp: 568-573.
- [5]Mohanty, S.P. and K.R. Ramakrishna, 2000. A DCT domain visible watermarking technique for images. Proceedings of the IEEE International Conference on Multimedia and Expo, July 30-Aug. 2, New York City, USA., pp: 1029-1032.
- [6] Tsai, M.J., 2009. A visible watermarking algorithm based on the content and contrast aware (COCOA) technique. *J. Vis. Commun Image R.*, 20: 323-338.
- [7] Yang, Y., X. Sun, H. Yang, C. Li and R. Xiao, 2009. A contrast-sensitive reversible visible image watermarking technique. *IEEE Trans. Circuits Syst. Video Technol.*, 19: 659-667.
- [8] Zeng, W., & Wu, Y. (2010). A visible watermarking scheme in spatial domain using HVS model. *Information Technology Journal*, 9(8), 1622-1628.
- [9] Khan, A., X. Niu and Z. Yong, 2008. A robust framework for protecting computation results of mobile agents. *Inform. Technol. J.*, 7: 24-31.
- [10] Yong, L., C. Li-Zhi, X. Zhi-Hong and W. Yi, 2004. A visible digital watermark based on integer wavelet transform with parameters. *J. Software*, 15: 238-249.

# DetECCIÓN DE LA COBERTURA DE MANGLAR MEDIANTE EL PROCESAMIENTO DE IMÁGENES SPOT-5

## Segmentación basada en NDVI y textura

Elizabeth Manríquez Duarte  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
elidium@gmail.com

Guillermo Martínez Flores  
Departamento de Oceanología  
Centro Interdisciplinario de Ciencias  
Marinas del IPN  
La Paz, Baja California Sur, México  
gmflores@ieee.org

Jesús Antonio Castro  
División de Estudios de Posgrado e  
Investigación  
Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México  
jcastro@itlp.edu.mx

**Resumen**—Se utilizan dos métodos de segmentación en imágenes de satélite multiespectrales SPOT-5 para obtener la segmentación de la cobertura de manglar en dos regiones al sur de la Bahía de La Paz, B. C. S. Los métodos utilizados fueron: segmentación basada en el NDVI y segmentación por textura basada en entropía.

**Palabras clave**—Segmentación, vegetación, NDVI, Textura, clasificación, entropía.

### I. INTRODUCCIÓN

El procesamiento de imágenes de satélite es utilizado en la actualidad para una variedad de aplicaciones, tales como agricultura, monitoreo de áreas de desastre, medición de las temperaturas terrestres y marinas, monitoreo de cobertura vegetal, generación de mapas de pesca más probable, entre otras.

La cobertura vegetal puede ser detectada por diferentes métodos. Uno de los más comunes es utilizando índices de vegetación, calculados a partir de operaciones entre bandas espectrales captadas por satélites de observación de la Tierra. En función de la respuesta espectral de estas bandas, el índice arrojará valores altos para regiones con vegetación y muy bajos para otros elementos de la escena, como agua, suelo y construcciones. El Índice de Vegetación de Diferencia Normalizada (NDVI por sus siglas en inglés), es el más utilizado, pues reduce el efecto de la degradación del sensor y la influencia de efectos atmosféricos, además de ser muy sencillo de calcular [1].

Por lo general es difícil distinguir entre diferentes tipos de vegetación utilizando sólo este tipo de índices, aunque esto depende en buena medida de otros factores como la resolución espacial de las bandas, las condiciones de iluminación, la topografía del terreno, etc.

Los manglares representan un tipo de vegetación transicional entre el ambiente marino y propiamente terrestre. La existencia y funcionalidad de los manglares es importante porque permite la diversidad de micro-ambientes y, en consecuencia, facilita la protección, crianza y alimentación de vida marina en general [2]. Es por ello que surge el interés de

implementar procedimientos que contribuyan a evaluar cambios en la cobertura de estos ambientes, utilizando imágenes de satélite captadas en diferentes épocas.

En este trabajo se propone utilizar, además del NDVI, un criterio de segmentación basado en textura, con la finalidad de dar más certidumbre al proceso de identificación de manglares.

Para llevar a cabo esta segmentación, se caracteriza la textura en la imagen por la uniformidad de intensidades en regiones locales, con base en el concepto de entropía. Normalmente los objetos en una imagen no mantienen una intensidad uniforme, pero pueden contener variaciones en las intensidades características de dicho objeto.

### II. DATOS Y ÁREA DE ESTUDIO

Para realizar este trabajo se utilizó una imagen SPOT-5 del 27/06/2010, proporcionada por ERMEX NG, proyecto conjunto entre la SAGARPA y la SEDENA. De esta imagen se extrajeron las bandas espectrales con resolución espacial de 10 m, que se muestran en la Tabla 1.

Tabla 1. Bandas Spot-5 utilizadas en este trabajo.

No.	Banda (nm)	Color
1	500-590	Verde
2	610-680	Rojo
3	780-890	Infrarrojo cercano
4	1,580-1,750	Infrarrojo de onda corta

Como regiones de prueba para los algoritmos, se seleccionaron dos zonas de manglar ubicadas al sur de la Bahía de La Paz, específicamente en la barrera arenosa El Mogote, así como en la Isla Espíritu Santo, Baja California Sur (Fig. 1).



Figura 1. Ubicación de las zonas de manglar seleccionadas para probar los algoritmos. a) Isla Espíritu Santo, b) barrera arenosa El Mogote, en la Bahía de La Paz, Baja California Sur.

### III. MÉTODOS

Se utilizaron dos métodos de segmentación para detectar la vegetación: segmentación basada en NDVI y segmentación por textura basada en entropía.

#### A. Segmentación basada en el NDVI

El NDVI es un índice usado para estimar la cantidad, calidad y desarrollo de la vegetación. Se calcula a partir del valor contenido en la misma posición de un pixel en la banda infrarroja cercana y la banda roja de una imagen ráster. El valor calculado es el resultado del valor en la capa infrarroja cercana menos el valor en la capa rojo visible, dividido entre el valor de la capa infrarroja cercana más el valor de rojo visible [3], como se muestra en (1).

$$NDVI = \frac{IRCercano - ROJO}{(IRCercano + ROJO)} \quad (1)$$

El cálculo del NDVI para un pixel dado da como resultado un número en el rango  $\pm 1$ . Los valores menores a 0 representan zonas sin vegetación, mientras que valores cercanos a 1 representan áreas con vegetación vigorosa [3].

#### B. Segmentación por textura basada en entropía

Un método para describir regiones consiste en cuantificar su contenido de textura. Este descriptor proporciona medidas de propiedades tales como suavizado, rugosidad y regularidad [4].

Existen tres tipos de métodos en el procesamiento de imágenes para describir la textura de una región. Estos son los estadísticos, los estructurales y los espectrales. Existen diferentes métodos estadísticos para describir la textura de una región, para este trabajo se utiliza un método estadístico basado en la entropía [5].

La entropía es un concepto relacionado con la incertidumbre sobre el valor de una variable aleatoria y el desorden de su función de distribución de probabilidad. Si la variable puede tomar una serie de valores con mayor probabilidad que otros, la incertidumbre de conocerla sería baja y su entropía también. Por el contrario, si todos sus posibles valores fuesen equiprobables, habría alta incertidumbre sobre su valor y la entropía sería alta [6]. La formulación utilizada para calcular la entropía  $H$  en una imagen  $X$  es (2).

$$H(X) = - \sum_{i=1}^n P(x_i) \log P(x_i) \quad (2)$$

Donde  $P(x_i)$  es la probabilidad de ocurrencia de  $x_i$ .

Una vez calculada la imagen de entropía, se generó una imagen binaria a partir de la especificación de un umbral en el valor de la entropía. Posteriormente se utilizaron operadores binarios para procesar la imagen resultante, generando una imagen binaria que delimita las regiones de manglar (Fig. 2).

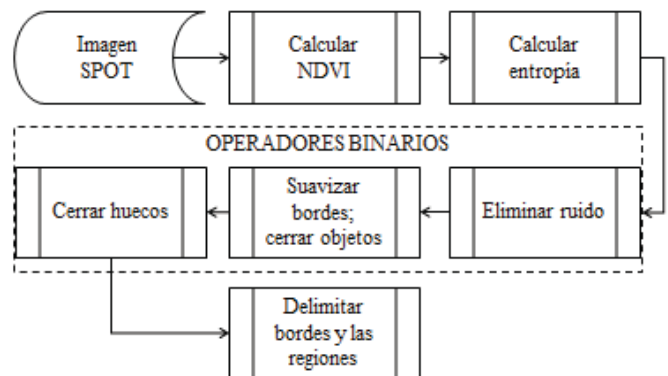


Figura 2. Diagrama de flujo del proceso de segmentación basado en entropía.

#### IV. RESULTADOS Y DISCUSIÓN

La imagen obtenida con el NDVI es calculada a partir de operaciones algebraicas entre las bandas espectrales correspondientes al rojo e infrarrojo. El resultado permite obtener una nueva imagen, en la que se destacan las regiones que presenta cobertura de vegetación (Fig. 3).

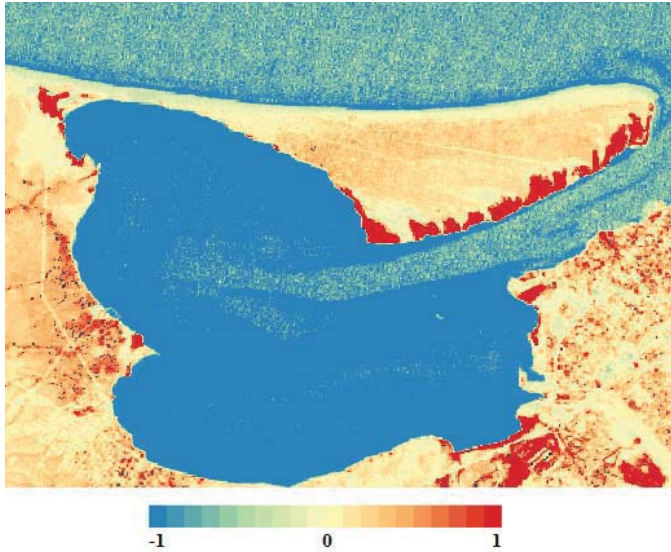


Figura 3. Resultado de la segmentación basada en el NDVI, para la barrera arenosa El Mogote. Las tonalidades en rojo (valores cercanos a 1) representan los parches de vegetación más vigorosa.

Los píxeles con valores cercanos a 1 representan la vegetación más vigorosa, mismos que aparecen en tonos rojizos. La zona marina se presenta, según esta escala de color, en tonos azules, el suelo sin vegetación y las zonas más someras de la región marina se muestran en amarillo. La vegetación menos vigorosa, como matorrales, aparece en tonos anaranjados más claros. En la imagen destacan diferentes puntos de vegetación en rojo, lo cual indica que se tiene que implementar un criterio de selección para descartar la vegetación que no sea manglar.

Con la segmentación por textura basada en entropía se pudo obtener una segmentación de la zona deseada (Fig. 6). Su base es calcular la entropía de los valores calculados con el NDVI. Tomando en cuenta una matriz de correlación para evaluar la vecindad de píxeles, se estableció un umbral de 0.85 en la imagen de entropía para obtener la imagen binaria.

En el resultado final se observa que se obtiene la segmentación solo de las zonas de manglar. En algunos parches segmentados se observa que se considera dentro de la segmentación pequeñas zonas que no son vegetación, lo cual se espera solucionar utilizando imágenes con mayor resolución espacial.

En las Figuras 4, 5 y 6 se muestra el resultado de la segmentación por textura basada en entropía para una región de la Isla Espíritu Santo, generado con datos provenientes de la ERMEXNG-0169 [2016].

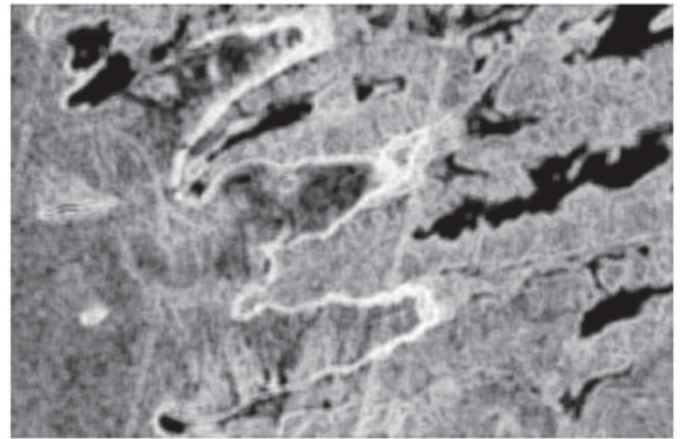


Figura 4. Representación de la entropía.



Figura 5. Imagen binaria resultado de la especificación de un umbral en la imagen de entropía.



Figura 6. Regiones segmentadas sobre la imagen en color verdadero.

## V. CONCLUSIONES

Con el método basado en NDVI fue posible detectar la cobertura de manglar. Sin embargo, no fue posible separar específicamente el manglar de otro tipo de vegetación. Con el método basado en textura, la delimitación de las áreas de manglar no fue precisa, pues se agruparon píxeles que corresponden a otro tipo de cobertura.

Como trabajo futuro se planea incorporar imágenes de mayor resolución espacial y mejorar los algoritmos, para tratar de obtener resultados más precisos. Con esto se podrán proponer procedimientos para el monitoreo de las zonas de manglar con imágenes multitemporales, cuantificando los cambios en su cobertura.

## AGRADECIMIENTOS

Agradecemos al proyecto ERMEX NG de la SAGARPA y la SEDENA, por haber proporcionado las imágenes de satélite SPOT-5 utilizadas, así como al Instituto Tecnológico de La Paz

y al Centro Interdisciplinario de Ciencias Marinas del Instituto Politécnico Nacional, por las facilidades para realizar este trabajo.

## REFERENCIAS

- [1] Pérez Gutiérrez C., Muñoz Nieto A. L., “Teledetección: Nociones y Aplicaciones”, Universidad de Salamanca, Julio 2006
- [2] Félix Pico E., Serviere Zaragoza E., Riosmena Rodríguez R., León de La Luz J. L., “Los Manglares de la Península de Baja California”, Publicación de divulgación del Centro de Investigaciones Biológicas del Noroeste, 2011
- [3] Díaz García-Cervigón J. J., “Estudio de Índices de vegetación a partir de imágenes aéreas tomadas desde UAS/RPAS y aplicaciones de estos a la agricultura de precisión”, Junio 2015
- [4] González R. C., Woods R. E., “Tratamiento digital de imágenes”, Addison-Wesley/Diaz de Santos. 1992.
- [5] Shannon C. E. A Mathematical Theory of Communication. Bell System Technical Journal, 1943, 27 (3):379–423.
- [6] Carrión Pérez P. A., Ródenas García J., Rieta Ibáñez J. J., “Ingeniería Biomédica: Imágenes Médicas”, Ediciones de la Universidad de Castilla-La Mancha, 2006.

# Implementación de un algoritmo para la detección de polvo en imágenes MODIS

Guillermo Martínez Flores

Departamento de Oceanología  
Centro Interdisciplinario de Ciencias  
Marinas del IPN  
La Paz, Baja California Sur, México  
gmflores@ieee.org

Yara García Álvarez

Instituto Tecnológico de La Paz  
La Paz, Baja California Sur, México,  
yara.garcia@itlp.edu.mx

José Antonio Segovia Zavala

Universidad Autónoma de Baja  
California, Instituto de  
Investigaciones Oceanológicas,  
Ensenada, Baja California, México  
jsegovia@uabc.edu.mx

**Resumen**—El objetivo de este trabajo fue implementar un algoritmo para detectar tormentas de polvo, utilizando las bandas visibles e infrarrojas captadas por el sensor MODIS. Adicionalmente, con la finalidad de optimizar el proceso de selección de imágenes, se desarrolló un procedimiento basado en el análisis de datos de viento para determinar las épocas con mayor probabilidad de generar tolvaneras en regiones, específicas, en este caso los desiertos arenosos que rodean el Golfo de California (GC). Los resultados indican que la región norte del área de estudio tiene mayor potencial de actuar como fuente de material terrígeno al GC, principalmente la costa norte de Sonora y la región del Desierto de Altar, durante los meses de invierno y primavera.

**Palabras clave**—Polvo; imágenes multiespectrales; MODIS.

## I. INTRODUCCIÓN

Los aerosoles son pequeñas partículas ( $< 100 \mu\text{m}$ ) sólidas o líquidas presentes en suspensión en la atmósfera. Además de ser generados por eventos naturales como las tormentas de arena o erupciones volcánicas, estos aerosoles son también emitidos en grandes cantidades por actividades antropogénicas en centros urbanos e industriales, quema de biomasa asociada con los procesos de uso de suelo, alcanzando niveles que afectan tanto al clima como a la salud humana [1], [2]. Una fuente importante de metales bioactivos (p. ej. Fe, Cu, Al, Cd), ya sea de origen natural o antropogénico a la superficie del mar es a través del transporte eólico de origen terrígeno, principalmente desde los grandes desiertos del mundo. Estos ambientes son particularmente sensibles a las presiones de cambio global, y tales cambios podrían alterar la productividad del océano y por lo tanto el clima. El hierro es un nutriente esencial para todos los organismos, es usado en una variedad de sistemas de enzimas, incluyendo los de la fotosíntesis, respiración y fijación del nitrógeno. Para el fitoplancton, que normalmente está alejado de los sedimentos del fondo marino (rico en hierro), los requerimientos fisiológicos de este mineral se pueden alcanzar a través del polvo que se deposita en la superficie marina. El abastecimiento de hierro es por lo tanto un factor limitante del crecimiento del fitoplancton sobre vastas áreas del océano [3].

Estudios recientes [4] sobre flujo atmosférico de hierro hacia el Golfo de California ( $7.0$  a  $90.4 \mu\text{mol m}^{-2} \text{d}^{-1}$ ), revelan

que éste es del mismo orden de magnitud que los reportados para el Mar Rojo ( $1.5$  a  $116 \mu\text{mol m}^{-2} \text{d}^{-1}$ ; [5]), que son de los más altos del mundo. La determinación de las fuentes de aporte de estos elementos tiene implicaciones importantes, tanto en aspectos de contaminación marina como de productividad primaria, por lo que su ubicación espacial y la evaluación de su capacidad de aporte son componentes relevantes en la interacción continente-atmósfera-océano.

Debido al reconocimiento de la importancia de las propiedades de los aerosoles para estudios sobre el clima, la salud humana y sobre su impacto en el ambiente, en las últimas tres décadas se han desarrollado una serie de sensores instalados a bordo de satélites, que permiten el análisis cualitativo y cuantitativo de sus propiedades ópticas, proveyendo información sobre su distribución global, la variación anual e interanual de sus flujos y la identificación de sus fuentes (polvo originado en regiones desérticas, de incendios forestales, regiones industriales, etc.) [1]. Actualmente la constelación de satélites A-Train (Afternoon-Train) incluye los satélites Aqua, CloudSat, CALIPSO (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) y Aura, que en conjunto incorporan 15 sensores para analizar varios rasgos de la atmósfera, océanos y suelos, proporcionando una visión completa sobre los cambios en el planeta. Los datos proporcionados por CALIPSO ofrecen una nueva visión sobre el papel que juegan el polvo, nubes y los aerosoles atmosféricos en el clima [6].

El objetivo general de este trabajo es implementar los algoritmos para la detección y visualización de polvo atmosférico en imágenes del noroeste de México captadas con el sensor MODIS, poniendo énfasis en el transporte de polvo hacia el GC.

## II. DATOS Y MÉTODOS

Las imágenes MODIS Nivel 1b (MYD021KM) fueron obtenidas del Sistema de Distribución y Archivo de Atmósfera Nivel 1 (LAADS, pos sus siglas en inglés), desde su sitio [ladsweb.nascom.nasa.gov](http://ladsweb.nascom.nasa.gov). Estos archivos contienen las bandas de radiancia y emisivas con resolución espacial estandarizada a  $1 \text{ km}$ . La Fig. 1 muestra las fases generales del pre-procesamiento de estas imágenes en un diagrama conceptual. Éstas consisten en: reprojectar y georreferenciar las imágenes,

que originalmente tienen proyección sinusoidal, (Fig. 2), extraer la región de interés, seleccionar las bandas necesarias para aplicar los algoritmos de detección de polvo e incluir la máscara binaria tierra/agua [7]. La máscara binaria se utiliza para aplicar dos tipos de algoritmos [8], [9]: detección de polvo sobre áreas marinas y detección sobre áreas continentales. El pre-procesamiento consiste en generar una imagen con ocho bandas, siete en el rango visible e infrarrojo del espectro electromagnético (bandas 1 a 4, 26, 31 y 32) y una correspondiente a la máscara binaria de tierra/agua. Estas imágenes de 8 bandas son el insumo para la detección de polvo, algoritmos que se ilustran en la Fig. 3. Los procedimientos descritos en este apartado fueron implementados en IDL y ENVI.

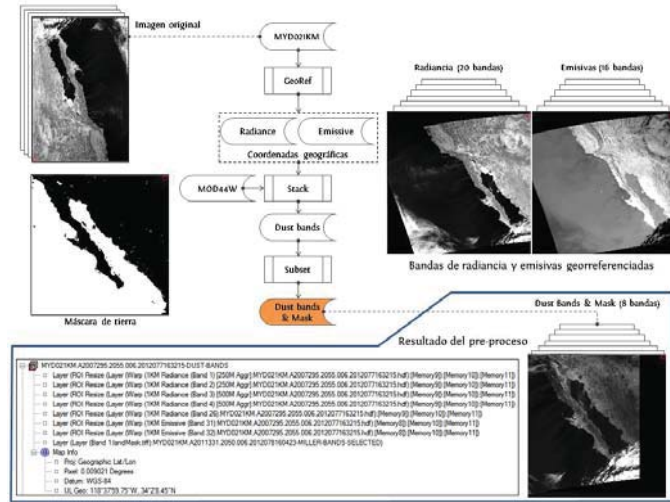


Figura 1. Pre-procesamiento de las imágenes MODIS Nivel 1b (MYD021KM). Se esquematiza la georeferenciación de las bandas de radiancia y emisivas, la extracción de la región de interés a partir de una máscara de tierra (MOD44W) y la selección de bandas necesarias para aplicar los algoritmos de detección de polvo sobre agua y sobre tierra [8], [9].

Con la finalidad de excluir píxeles cercanos a la costa, se aplicó el operador morfológico de erosión (Fig. 2) a la máscara binaria [10] antes del proceso de detección de polvo. Este procedimiento permite eliminar la interferencia o ruido causado (en los algoritmos de detección de polvo) por el oleaje cercano a la costa, así como por la reflectancia en aguas someras. El efecto de este procedimiento se ve reflejado en el resultado como una franja sin datos en torno a la línea de costa. De no considerar esta exclusión de píxeles, se podrían generar falsas detecciones de polvo en esta franja.

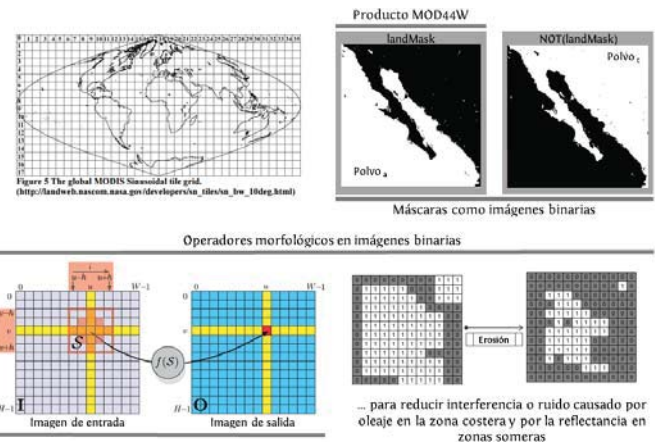


Figura 2. Arriba a la izquierda, malla global para las imágenes MODIS con proyección sinusoidal. Arriba a la derecha, máscaras binarias tierra/agua (producto MOD44W). La parte inferior de la figura esquematiza la aplicación del operador morfológico de erosión sobre la máscara binaria para excluir píxeles cercanos a la costa [10].

Una vez concluida la etapa de pre-procesamiento, se aplicaron los algoritmos para la detección de polvo. Para ello se consideran dos criterios [8]: 1) la proporción entre la banda roja y azul para diferenciar el polvo de las nubes; y 2) la diferencia de temperatura entre la superficie continental y las regiones con polvo para detectar polvo sobre tierra. Este procedimiento se ilustra en la Fig. 3, a la izquierda se muestra un diagrama de bloques indicando las longitudes de onda centrales de cada banda implicada en ambos algoritmos. Los resultados de cada uno (detección de polvo sobre mar y sobre tierra), se asignan al cañón rojo del video, los datos de la banda 4 al cañón verde y los de la banda 3 al cañón azul, para desplegar las composiciones en falso color ( $P_a$  y  $P_t$ , Fig. 3). Finalmente esas dos composiciones son fusionadas para obtener la imagen resultado ( $P_{at}$ ).

A la derecha de la Fig. 3, se muestra el detalle del algoritmo propuesto por Miller [8] y revisado por Chen et al., [9]. El procedimiento descrito en el diagrama de flujo del algoritmo resuelve la detección de polvo sobre tierra (normalmente ésta es más difícil que la realizada sobre regiones marinas, pues el contraste es mayor en regiones donde aparece el polvo sobre mar). Sólo la última línea del algoritmo corresponde a la detección de polvo sobre cuerpos de agua (utilizando el criterio de proporción entre los canales rojo y azul). En este diagrama, la función modisBright convierte los valores de emisividad de las bandas 31 y 32 a valores de temperatura en grados Kelvin, la función Normalize ajusta los valores de la variable especificada en el rango definido por el segundo parámetro de la función.



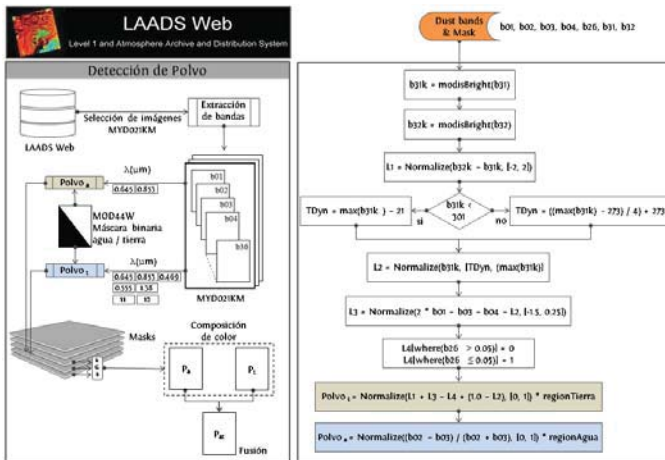


Figura 3. A la izquierda, diagrama de bloques indicando las longitudes de onda centrales de las bandas utilizadas en cada algoritmo (detección de polvo sobre agua [Pa] y sobre tierra [Pt]). Los resultados de éstos son asignados al cañón rojo de video, (al cañón verde se asigna la banda 4 y al azul la banda 3) generando las composiciones en falso color RGB  $P_a$  y  $P_t$  para fusionarlos posteriormente en la imagen  $P_a$ . A la derecha, diagrama de flujo que detalla los algoritmos para la detección de polvo propuestos por Miller [8] y revisados por Chen et al. [9].

Adicionalmente, con la finalidad de optimizar el proceso de selección de imágenes MODIS, se desarrolló un procedimiento basado en el análisis de datos de viento y en las condiciones del terreno para determinar las épocas y regiones con mayor probabilidad de generar tolvaneras en regiones aledañas al GC. Los datos de viento fueron obtenidos de la base de datos del Reanálisis Regional de América del Norte (NARR, por sus siglas en inglés) de la NOAA ([www.esrl.noaa.gov](http://www.esrl.noaa.gov)). Las regiones fueron seleccionadas del Sistema Zonas Áridas del Servicio de Información Agroalimentaria y Pesquera (SIAP), dependiente de la Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). Para llevar a cabo esta evaluación, se generaron mapas de dirección e intensidad del viento anuales con resolución temporal de 3 horas y espacial de 0.3 grados (32 km), sobre las regiones susceptibles a erosión eólica.

### III. RESULTADOS

La Fig. 4 muestra el resultado de los procedimientos descritos en el apartado anterior, aplicados a una imagen del 27/11/2011, en la que se observan plumas de polvo originadas en los desiertos del Vizcaíno y en la región costera de Sonora. Esta última cruza el Golfo de California, trasladándose de 100 km (en el trayecto más corto) hasta más de 200 km, al sur de Santa Rosalía, Baja California Sur. La porción más densa de la pluma que se origina en el desierto del Vizcaíno se traslada a aproximadamente 250 km desde su origen en el centro de la península hacia el Océano Pacífico. El patrón de vientos hacia el oeste que da lugar al transporte de polvo en esta región no es común, pues la época de invierno se caracteriza por una dominancia de los vientos del noroeste.

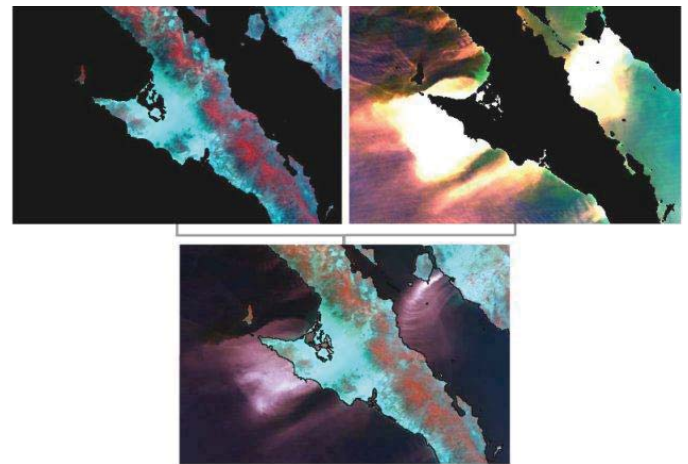


Figura 4. Imagen del 27/11/2011. Arriba los resultados de los algoritmos para la detección de polvo sobre tierra y mar. Abajo se muestra el resultado de la fusión de imágenes. En ésta, la franja en color negro en torno a la línea de costa corresponde a la región descartada por el operador morfológico descrito en la Fig. 2.

En la imagen del 22/10/2007 (Fig. 5) se resaltan múltiples plumas de polvo originadas en la costa oriental de la Península de Baja California, desplazándose a distancias mayores a los 500 km hacia el Océano Pacífico. Al norte, en la región cercana a San Diego, E. U., las plumas de humo están asociadas a incendios forestales. Desde el Desierto de Sonora se originan tolvaneras que transportan polvo hacia el sur de Las Grandes Islas, recorriendo aproximadamente 350 km y llegando hasta las costas de la península.

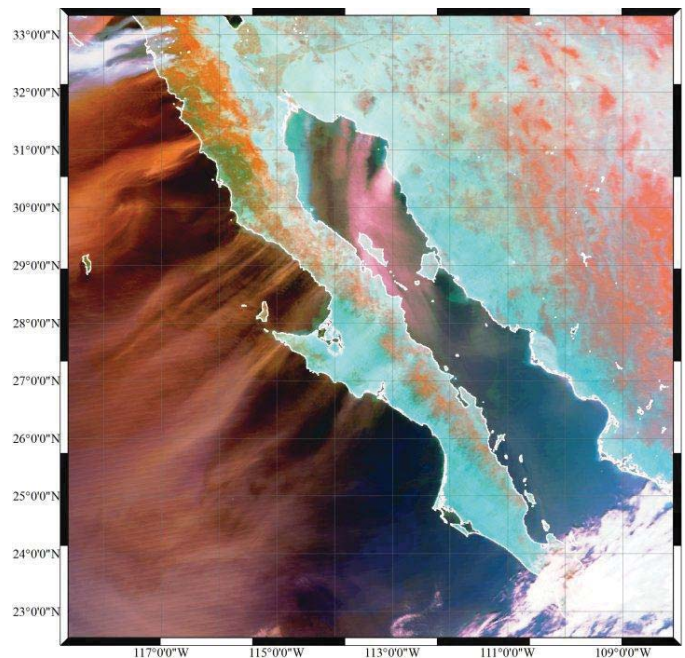


Figura 5. Imagen del 22/10/2007. Múltiples plumas de polvo transportadas desde la Península de Baja California hacia el Océano Pacífico y desde el Desierto de Sonora hacia la región norte del GC. Las plumas blancas en el área de San Diego, E. U. corresponden a humo asociado a incendios forestales.

Los resultados de la caracterización del terreno dan una primera aproximación a la identificación de regiones potencialmente susceptibles a aportar aerosoles por transporte

oólico hacia el GC. Estas regiones se localizan principalmente en la región norte del golfo (Desierto de Altar, Desierto de Sonora, Este de la Sierra San Pedro Mártir). En la región sur de la península, con vientos soplando al Este, las posibles fuentes de material terrígeno al Golfo de California son el Desierto del Vizcaíno y los Llanos de Magdalena. Las condiciones del viento más favorables (rachas transversales al golfo) para este aporte al GC se presentan durante invierno y primavera.

#### IV. CONCLUSIONES

Los algoritmos implementados resultaron adecuados para caracterizar el transporte de polvo que se origina en el noroeste de México, sin embargo se sugiere implementar mejoras al contraste de la imagen o al algoritmo de detección de polvo sobre tierra para poder inferir con mayor exactitud las áreas en las que se originan las tolvaneras.

El aporte polvo hacia el GC ocurre con mayor intensidad en su región norte durante los meses de invierno y primavera, principalmente desde los Desiertos de Altar y de Sonora. Los vientos transversales al golfo, que favorecen el aporte de polvo hacia el GC, no son una condición común en la región, pues en general, en ésta predominan los vientos del noroeste en invierno y del sureste en verano, por lo que estos patrones transversales podrían considerarse como rachas intensas ( $> 6 \text{ m}^{-5}$ ) de periodos cortos, por lo que se propone realizar trabajos sobre la persistencia del viento en la región que permitan llegar a conclusiones en este sentido.

#### AGRADECIMIENTOS

Este trabajo fue realizado como parte del proyecto CONACYT CB-2011-01-166897 “Flujo atmosférico de metales bioactivos y su solubilidad en el Golfo de California: un escenario hacia el cambio climático”. Los datos MODIS utilizados en este estudio fueron obtenidos de la División Sistema Tierra-Sol de la NASA, archivados y distribuidos por el Sistema de Procesamiento Adaptativo de MODIS (MODAPS), desde su sitio [ladsweb.nascom.nasa.gov](http://ladsweb.nascom.nasa.gov). La máscara de tierra/agua fue obtenida del trabajo de Carroll et al., [7] (<http://glcf.umd.edu/data/watermask/>). Los datos de viento NARR fueron proporcionados por NOAA-ESRL Physical

Sciences Division, Boulder Colorado ([www.esrl.noaa.gov/psd/](http://www.esrl.noaa.gov/psd/)). La cartografía de zonas áridas fue proporcionada por el Sistema Zonas Áridas del Servicio de Información Agroalimentaria y Pesquera (SIAP, México). Los autores agradecen también al Instituto Tecnológico de La Paz y al Centro Interdisciplinario de Ciencias Marinas del Instituto Politécnico Nacional, por las facilidades prestadas para realizar este trabajo.

#### REFERENCIAS

- [1] King M., Kaufman Y., Tanré D. and Nakajima T. 1999. Remote sensing of tropospheric aerosols from space: Past, present, and future. *Bull Am Meteorol Soc*, 80: 2229-2259.
- [2] Arellano, H. J. 2011. Evaluación de aerosoles troposféricos del AMCM por medio de imágenes de satélite GOES. Tesis de Maestría. Centro Interdisciplinario de Investigaciones y Estudios sobre Medio Ambiente y Desarrollo (CIEMAD-IPN), México.
- [3] Jickells, T.D., Z. S. An, K. K. Andersen, A. R. Baker, G. Bergametti, N. Brooks, J. J. Cao, P. W. Boyd, R. A. Duce, K. A. Hunter, H. Kawahata, N. Kubilay, J. laRoche, P. S. Liss, N. Mahowald, J. M. Prospero, A. J. Ridgwell, I. Tegen, R. Torres. 2005. Global iron connections between dust, ocean biogeochemistry and climate. *Science*, 308, 67-71.
- [4] Segovia-Zavala, J. A., Delgadillo-Hinojosa, F., Lares-Reyes, M. L., Huerta-Díaz, M. A. 2009. Atmospheric iron input and the sea-surface dissolved iron concentrations in the Gulf of California. *Ciencias Marinas* 35:75-90.
- [5] Chase, Z., A. Paytan., K. S. Johnson., J. Street and Y. Chen. 2006. Input and cycling of iron in the Gulf of Aqaba, Red Sea. *Global Biogeochemical Cycles*, 20, GB3017, doi:10.1029/2005GB002646.
- [6] L'Ecuyer, T. S. and Jiang, J. H. 2010. Touring the atmosphere aboard the A-Train. *Physics Today*, 63(7), 36-41.
- [7] Carroll, M., Townshend, J., DiMiceli, C., Noojipady, P., Sohlberg, R. 2009. A New Global Raster Water Mask at 250 Meter Resolution. *International Journal of Digital Earth*. 2(4).
- [8] Miller, S. D. 2003. A consolidated technique for enhancing desert dust storms with MODIS. *Geophys. Res. Lett.*, 30, 2071, doi:10.1029/2003GL018279, 20.
- [9] Chen B., Zhang P., Zhang B., Jia R., Zhang Z., Wang T., Zhou T. 2014. An Overview of Passive and Active Dust Detection Methods Using Satellite Measurements. *J. Meteor. Res.*, 28(6): 1029-1040, 0, doi: 10.1007/s13 351-014-4032-4.
- [10] Corke, P. 2015. *Robotics, Vision and Control: Fundamental Algorithms in MATLAB*. Springer-Verlag Berlin Heidelberg, ISBN 978-3-642-20143-1. 570 p.

# Construcción de modelo acústico en español de México para el área pediátrica y su implementación con CMU-SPHINX.

Ing. Mario E. Villalobos Gallegos, M. C. Beatriz A. Olivares Zepahua  
 División de Estudios de Posgrado e Investigación  
 Instituto Tecnológico de Orizaba  
 Orizaba, Veracruz, México  
 villalobosm@acm.org, bolivares@ito-depi.edu.mx

Dr. Abel Herrera Camacho  
 Laboratorio de Tecnologías del Lenguaje  
 Universidad Nacional Autónoma de México  
 Ciudad de México, México  
 abelherrerac1@gmail.com

**Resumen—** Entre los usos que tiene el reconocimiento de voz se encuentra el ingreso de texto a documentos sin necesidad de utilizar el teclado de una computadora o dispositivo móvil. Además de ser usado en diversas áreas, tiene especial uso en el campo médico, por lo que se creó un modelo acústico que incluye la terminología de pediatría para posteriormente ser implementado en un hospital de la región de Orizaba, Veracruz. En este documento se presentan los avances alcanzados hasta el momento en la construcción del modelo acústico que incluya terminología del área pediátrica, así como pruebas de reconocimiento realizadas con CMU-SPHINX versión 4.

**Palabras clave—** modelo acústico; reconocimiento de voz

## I. INTRODUCCIÓN

La actualización del expediente clínico no es del agrado de muchos médicos y debido al poco tiempo con que cuentan para esta tarea la ejecución del mismo no es grato [1]. Entre las tareas que incluye la actualización del expediente se destaca la creación de notas médicas. Estas son de vital importancia para el seguimiento del historial de los pacientes que acuden a solicitar los servicios médicos de los nosocomios [2]. Se pretende construir una herramienta que haga uso del reconocimiento de voz para hacer más atractiva el ingreso de las notas médicas para los doctores como se ha demostrado en otros estudios similares [3,4].

Los sistemas de reconocimiento de voz pueden dividirse en dos tipos, dependientes del hablante e independientes del hablante. El primer tipo requiere de entrenamiento por parte del usuario final, mientras que el segundo no, además en el último se realiza un tipo de adaptación a la voz del usuario para mejorar el rendimiento [5].

Para que un sistema de reconocimiento de voz funcione es necesario contar con:

- 1) un diccionario fonético que contenga los términos que se buscan sean reconocidos,
- 2) archivos de audio con oraciones dictadas por las personas que utilizaran el sistema (para el caso de un sistema dependiente del hablante) y

3) la transcripción a texto de cada uno de los archivos de audio para ser identificadas en el proceso de entrenamiento.

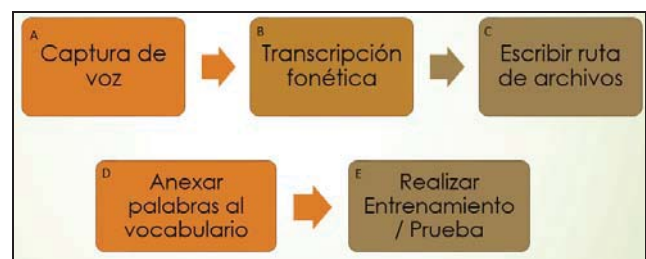
Una vez que se cuenta con lo anterior es posible realizar el entrenamiento del sistema de reconocimiento de voz, el cual en este caso es CMU-SPHINX versión 3 (o simplemente SPHINX) siguiendo los lineamientos descritos [6]. Posteriormente el modelo acústico resultante se usará con la versión 4 para Java ya que la aplicación final se implementará en esta plataforma.

El módulo de reconocimiento de voz aquí presentado forma parte de un sistema más amplio que se encuentra en construcción y hace uso de bibliotecas y herramientas de software libre para minimizar la inversión necesaria para el desarrollo e implementación en algún hospital de la región de Orizaba, Veracruz.

## II. METODOLOGÍA

Un modelo acústico contiene propiedades acústicas que incluyen vectores de características para cada fonema o unidad sonora. Para el entrenamiento de dicho modelo acústico se tomó como base el corpus del Proyecto CIEMPIESS (Corpus de Investigación en Español de México del Posgrado de Ingeniería Eléctrica y Servicio Social) [7]

Fig. 1. Pasos para entrenar modelo acústico.



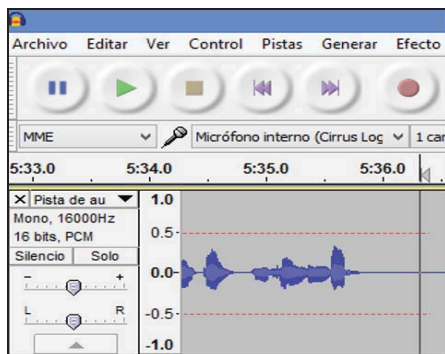
El CIEMPIESS consta de 17 horas de audio recopilados de programas de radio. A estas horas de audio se agregó más tiempo para realizar el entrenamiento del sistema con los términos médicos y así aumentar el porcentaje de acierto para este caso.

En la Fig. 1 se muestra de manera general el proceso realizado para el entrenamiento del modelo acústico y más adelante se detalla el mismo.

#### A. Captura de voz

Para este fin se utilizó el software de grabación y edición de audio Audacity desarrollado por Audacity Team [8]. Se instala y se procede a la configuración del entorno antes de iniciar la grabación tal como se indica en [9]. En las preferencias debe indicarse la frecuencia de muestreo a 16,000Hz, el formato de muestra como 16-bit y cambiar los canales de grabación a 1(Mono).

Fig. 2. Niveles de grabación máximos recomendados.



Es muy importante no rebasar los niveles indicados en la Fig. 2, esto para no generar ruidos. Una vez grabado el audio es necesario elegir porciones de este con una duración de entre 5 y 30 segundos aproximadamente para posteriormente exportarlos en formato WAV de Microsoft PCM de 16 bit con signo, ya que este es el formato admitido por SPHINX para realizar el entrenamiento [6]. Cada archivo exportado contiene lecturas de los enunciados contenidos en el Manual Harriet Lane de Pediatría [10], específicamente del capítulo 20 que se refiere a la neurología pediátrica. En total se tienen 108 archivos de audio con la terminología médica.

#### B. Transcripción fonética

A diferencia de la ortografía, la transcripción fonética no expresa significado lingüístico-semántico; lo que hace es expresar un análisis de la pronunciación, por ejemplo haya (árbol) y halla (dar con alguien que se busca) se pronuncian de la misma manera [11] y la transcripción fonética para ambos casos es la misma. Dicho en otras palabras, es la manera en cómo se pronuncian las palabras, no cómo se escriben. Es importante realizarla para la construcción del diccionario fonético explicado más adelante en la sección D, para ello debe efectuarse un proceso de pre-transcripción [7] usando el

alfabeto fonético Mexbet [12]. En el proyecto CIEMPIESS se incluye una función codificada en Python para encontrar la vocal tónica, pero se realiza para una sola palabra, por lo que fue necesario escribir una pequeña rutina que permitiera agilizar el proceso y realizar la transcripción de todo el texto de una sola vez, dicha rutina fue codificada con el entorno de desarrollo LiClipse [13]. En el Fig. 3 se muestra esta rutina.

Fig. 3. Código para encontrar la vocal tónica.

```

1  from vocal_tonica import vocal_tonica
2  f = open('entrada.txt','r')
3  r = open('resultado.txt','w')
4  i = 73
5  for linea in f:
6      palabras = linea.split(' ')
7      linea_nueva = "<s><sil>"
8      for pal in palabras:
9          nueva_pal = vocal_tonica(pal)
10         linea_nueva = linea_nueva + nueva_pal + " "
11         linea_nueva = linea_nueva + "<sil></s>"+("mario00" + str(i) + ") \n"
12         i = i + 1
13         r.write(linea_nueva)

```

Cabe destacar que en la línea 11 se inserta el nombre del archivo de audio, el cual es consecutivo para facilitar el proceso de identificación.

Al terminar lo anterior se procede a sustituir en el archivo resultante las letras ñ para evitar problemas de codificación y las letras x que en el español de México tienen 4 sonidos distintos según la palabra en donde se encuentren siguiendo los lineamientos indicados [7]. Como consecuencia en este paso se tiene un archivo con la pre-transcripción de todos los enunciados contenidos en los audios grabados y descritos en el apartado A.

#### C. Escribir ruta de archivos

Este paso se trata de indicar la ubicación física de cada uno de los archivos de audio para que el entrenador de SPHINX pueda localizarlos. De igual manera se indican los archivos con los que se realizará la ronda de prueba en línea de comando después de terminar el entrenamiento (en caso de ser necesaria este tipo de prueba).

#### D. Anexas palabras al vocabulario

Para que el sistema reconozca las palabras incluidas en el archivo de pre-transcripción es necesario incluirlas en el diccionario fonético o vocabulario, este es una lista de los términos que tiene el siguiente formato, se inicia con la pre-transcripción de la nueva palabra y seguido a esta la transcripción fonética de cada letra que la conforman separadas por un espacio, en la Fig. 4 se muestran algunos ejemplos.

Fig. 4. Ejemplos de transcripción fonética.

```

$antocromIIa s a n t o k r ( o m i _ 7 a
fiEEbre f i e _ 7 b r ( e
fibrOosis f i b r ( o _ 7 s i s
miocardIItis m i o k a r ( d i _ 7 t i s
mitocondriAAl m i t o k o n d r ( i a _ 7 l

```

A continuación, se explica a detalle la transcripción de la palabra xantocromía (coloración amarilla de una parte del cuerpo), en donde la “x” se reemplaza con \$ (signo de moneda) pues es el símbolo para el sonido /s/ el cual toma la letra “x” en este caso, además puesto que el sistema SPHINX no identifica mayúsculas de minúsculas [14] es necesario señalar la vocal tónica duplicando la vocal correspondiente, para este ejemplo el acento lo tiene la letra “i” quedando como “II” (doble i, en mayúsculas). Del lado derecho de la palabra escrita en su formato de pre-transcripción se define la transcripción fonética de la misma, en donde se observa que el signo \$ pasa a convertirse en “s”, la letra “c” pasa a convertirse en “k”, la letra “r” pasa a convertirse en “r(” y por último la vocal tónica “i” pasa a indicarse como “i\_7”. Esta transcripción se encuentra realizada tomando en consideración el alfabeto fonético Mexbet T22. Al finalizar este paso se cuenta con el diccionario fonético con los términos que están permitidos usar en el reconocedor de voz.

Por último, es necesario cambiar los archivos de audio al formato SPH (Sphere) para coincidir con el formato en el que se encuentra el corpus CIEMPIESS. Para esto se utilizó la aplicación en línea de comando SoX [15], la cual es una utilidad para cambiar el formato de archivos de audio. A continuación, se muestra la rutina usada para realizar la conversión de todos los archivos en una carpeta.

```
for i in /*.wav; do sox $i $.sph; done
```

#### E. Realizar Entrenamiento/Prueba

Cuando ya se cuenta con el vocabulario actualizado, los audios en formato SPH y el archivo de transcripción, se procede a realizar una ronda de entrenamiento, la cual consiste en convertir los archivos de audio en vectores de características, pues el sistema no trabaja directamente con las señales acústicas, lo anterior lo realiza el sistema SPHINX [6].

Cuando se ha completado el entrenamiento se cuenta con el modelo acústico y es posible realizar una ronda de prueba en línea de comando o incluir el modelo acústico entrenado en una aplicación mediante la creación de un archivo JAR en este caso pues se usa el entorno Java para implementar la solución.

Para realizar la decodificación en vivo y aumentar el porcentaje de acierto es necesario incluir una gramática, la cual es la manera en que el reconocedor espera se presentaran las palabras dictadas por el usuario. En la Fig. 5 se muestra una porción de la gramática utilizada.

Fig. 5. Porción del archivo de gramática.

```

#JSGF V1.0;

/**
 * JSGF Grammar
 */

grammar ES-MX;

public <descripcion> = [<arti>]<paci><anios>;
public <problema> = padEEce <padecimiento>;
public <subjetoivo> = sEE refiEEre cOOn <sintomas>;
//public <objetivo1> = sIIgnos vitAAles <signos>;
public <objetivo2> = <signos>;

```

Como se observa la gramática es similar a la notación de Backus-Naur, pero en este caso los enunciados válidos para ser reconocidos se definen como públicos y a continuación se especifica su estructura. Los términos que no son especificados como públicos, p. ej. <padecimiento>, solo pueden ser usados si se encuentran dentro de otro término que sí lo sea. Los términos entre corchetes se refieren a que son opcionales, es decir el usuario puede o no pronunciarlos. Para el caso del signo más (+) se refiere a que el término aparece una o más veces en el enunciado, mientras que el asterisco (\*) significa que la palabra se espera cero o más veces.

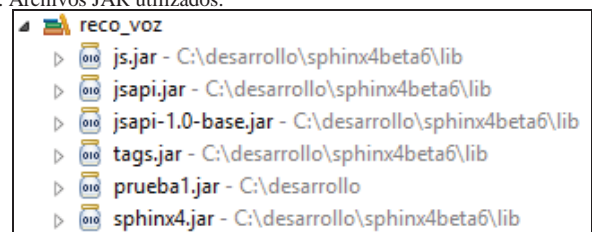
Cabe mencionar que tanto la conversión de formato con SoX, el proceso de entrenamiento del modelo acústico y la ronda de prueba en línea de comando fueron realizados en el sistema operativo (SO) Ubuntu Linux versión 12, pues se presentaron problemas al tratar de ejecutarlos en la versión 14 de dicho SO, así como para el caso de Windows 8 que es el SO en el que se desarrolló el módulo de reconocimiento de voz y se realizaron las pruebas de decodificación en vivo.

Otro detalle muy particular es el hecho de que se usó SPHINX versión 3 para crear el modelo acústico siguiendo el tutorial de [6], el cual esta codificado en el lenguaje C. Sin embargo, no hubo problemas para copiar el modelo acústico a la versión de Java pues utilizan los mismos tipos de archivo.

### III. RESULTADOS

El módulo de reconocimiento de voz se desarrolló en el lenguaje Java utilizando Eclipse Luna como ambiente de desarrollo. En la Fig. 6 se muestran las bibliotecas empleadas.

Fig. 6. Archivos JAR utilizados.



El archivo prueba1.jar cuenta con el modelo acústico entrenado y el diccionario de palabras. Los demás archivos son indispensables para el correcto funcionamiento de la biblioteca de SPHINX. A continuación, en el Fig. 7 se muestra parte del código resultante.

Fig. 7. Porción de código del módulo.

```

1 public String reconocer(){
2     String texto = "";
3     cm = new
      ConfigurationManager(Reconocimiento.class.getResource("config.xml"));
4     recognizer = (Recognizer) cm.lookup("recognizer");
5     microphone = (Microphone) cm.lookup("microphone");
6     recognizer.allocate(); microphone.initialize();
7     if (!microphone.startRecording()) {
8         recognizer.deallocate(); System.exit(0); }
9     Result result = recognizer.recognize();
10    if (result != null) texto = texto + "-" +
      convertirATextoNormal(result.getBestFinalResultNoFiller());
11    microphone.stopRecording(); recognizer.deallocate();
12    microphone.clear();
13    return texto; }
  
```

Hasta el momento se han agregado 133 términos del área médica que en conjunto con los incluidos en el diccionario del CIEMPIESS suman 53315 palabras.

Como se menciona en el inciso E de la sección anterior, es posible realizar una ronda de prueba en línea de comando, esto para obtener resultados preliminares de manera rápida. En la tabla 1 se muestran la tasa de error por oración obtenida en primer lugar con el modelo entrenado solo con los audios incluidos con el CIEMPIESS, en segundo lugar, con el audio que contiene la terminología médica.

TABLA 1. RESULTADOS EN LÍNEA DE COMANDO.

	Con audios del CIEMPIESS	Audios del CIEMPIESS + terminología médica
<b>Tasa de error</b>	59.2%	47.9%

Como se observa con agregar unos pocos minutos (20 minutos) resultó en una disminución del 11.3% en la tasa de error por oración.

Para el caso del reconocimiento en vivo los resultados también fueron favorecedores, la tabla 2 los muestra el porcentaje de acierto obtenido.

TABLA 2. RESULTADOS DE DECODIFICACIÓN EN VIVO.

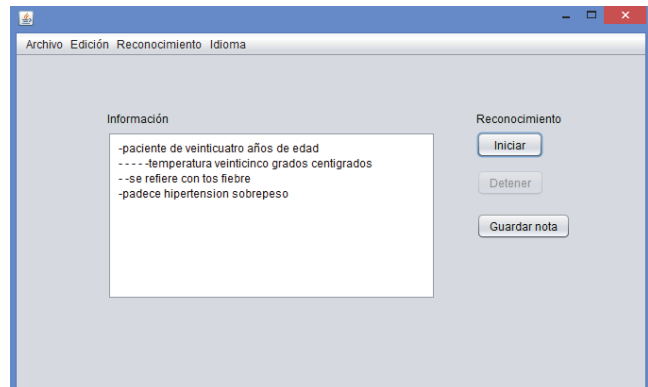
	Con audios del CIEMPIESS	Audios del CIEMPIESS + terminología médica
<b>Porcentaje de acierto</b>	62.25%	74.38%

Para los resultados mostrados en la tabla 2 se aumentó el porcentaje de precisión en un 12.13%, esto se logró con el uso de una gramática como la que se mostró en la Fig. 5 y a continuación se listan algunos ejemplos de los enunciados permitidos en donde, para el caso de los números pueden ser cualquiera entre el 1 y el 100 inclusive.

- El/La paciente de 23 años de edad.
- Padece diabetes, hipertensión, sobrepeso.
- Se refiere con tos, ojos rojos, fiebre.
- Temperatura 28 grados centígrados.

Por último, en la Fig. 8 se muestra la interfaz del prototipo con los términos escritos únicamente mediante reconocimiento de voz, aunque se cuenta también con la opción de corregir el texto mediante el teclado cuando así se requiera.

Fig. 8. Interfaz del prototipo.



#### IV. TRABAJOS A FUTURO

Se tiene planeado incluir más horas de audio para el entrenamiento del modelo acústico, además de crear un sistema independiente del usuario, lo que implica capturar la voz de alrededor de 200 personas [14], se aumentará el número de términos médicos que se reconocen y se actualizará la gramática para reflejar similitud con las notas médicas usadas en los expedientes clínicos.

Como se mencionó al inicio, el módulo aquí presentado es parte de una aplicación más grande la que permitirá el registro de notas médicas en el expediente clínico electrónico de un sanatorio de la región, actualmente se está trabajando en el desarrollo de esta aplicación para en meses posteriores realizar pruebas con los médicos del hospital.

Se busca alcanzar por lo menos un 90% de efectividad en pruebas en vivo, ya que se pretende agilizar la creación y actualización del expediente clínico mencionado en el párrafo anterior.

## V. CONCLUSIONES

Se encontró que al agregar más grabaciones de audio al entrenamiento se mejoraron los resultados del reconocimiento tanto en pruebas en línea de comando como en decodificación en vivo. Con lo anterior se infiere que al agregar más horas de audio se podrá obtener un mayor porcentaje de acierto, esto hará más atractivo el uso del sistema para el usuario final.

Se observó que el proceso de creación y actualización del modelo acústico para ser usado en un software de reconocimiento de voz es una tarea laboriosa y que requiere de mucho tiempo para ser realizada.

Una vez obtenido el modelo entrenado es muy fácil integrarlo en una aplicación, solo es necesario tener mucho cuidado en el archivo de configuración para indicar la ubicación de los archivos necesarios, como el diccionario de términos y el mismo modelo acústico.

## VI. AGRADECIMIENTOS

El primer autor agradece el apoyo económico de la beca número 329209 otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACYT), además se agradece al Tecnológico Nacional de México por el apoyo en el desarrollo de este proyecto y al Laboratorio de Tecnologías del Lenguaje de la Facultad de Ingeniería de la Universidad Nacional Autónoma de México (UNAM) por las facilidades otorgadas para la adquisición de conocimientos en el área de reconocimiento de voz, así como para el uso del corpus CIEMPIESS.

El segundo autor agradece al proyecto DGAPA-UNAM-PAPIIT IT 102314.

## REFERENCIAS

- [1] A. Shuman, "Speed EHR documentation with voice recognition software", *Contemporary Pediatrics*, Junio 2014. [En línea]. Disponible:

<http://contemporarypediatrics.modernmedicine.com/contemporary-pediatrics/content/tags/contemporary-pediatrics-peds-v20/speed-ehr-documentation-voice-?page=full>. [Último acceso: 26 enero 2016].

- [2] N. B. Delgado, "Comisión nacional de arbitraje medico", 3 febrero 2016. [En línea]. Disponible: [http://www.conamed.gob.mx/comisiones\\_estatales/coesamed\\_nayarit/publicaciones/pdf/expediente\\_clinico.pdf](http://www.conamed.gob.mx/comisiones_estatales/coesamed_nayarit/publicaciones/pdf/expediente_clinico.pdf).
- [3] M. Koivikko, "Improvement of report workflow and productivity using speech recognition: A follow-up study", *Journal of digital imaging*, vol. 21, n° 4, pp. 378-382. DOI: 10.1007/s10278-008-9121-4, 2008.
- [4] W. y. F. W. Xinxin, "The application of speech recognition in radiology information system", *International Conference on Biomedical Engineering and Computer Sciences*, pp. 1-3. DOI:10.1109/ICBECS.2010.5462425, 2010.
- [5] H. y. L. K. Xuedong, "On speaker-independent, speaker-dependent, and speaker-adaptive speech recognition", *IEEE Transactions on speech and audio processing*, vol. 1, n° 2, pp. 150-157, 1993 DOI:10.1109/89.222875.
- [6] Carnegie Mellon University, "Robust Group Tutorial", [En línea]. Disponible: <http://www.speech.cs.cmu.edu/sphinx/tutorial.html>. [Último acceso: 16 noviembre 2015]
- [7] J. Hernández-Mena y A. Herrera-Camacho, "Proyecto CIEMPIESS-UNAM", UNAM, [En línea]. Disponible: <http://www.ciempiess.org/about>. [Último acceso: 16 noviembre 2015].
- [8] A. Team, "Audacity", [En línea]. Disponible: <http://audacityteam.org/?lang=es>. [Último acceso: 16 noviembre 2015].
- [9] VoxForge, "voxforge.org", [En línea]. Disponible: <http://www.voxforge.org/home/dev/acousticmodels/linux/create/htkjulius/tutorial/data-prep>. [Último acceso: 2 febrero 2015].
- [10] M. T. y. K. Arcara, "Manual Harriet Lane de Pediatría", Barcelona: Elsevier España, 2013.
- [11] B. Heselwood, "Phonetic transcription in theory and practice", Edinburgh: Edinburgh University Press Ltd, 2013.
- [12] J. Cuétara-Priede, "Fonética de la ciudad de México. Aportaciones desde las tecnologías del habla", México: Universidad Nacional Autónoma de México, tesis de maestría 2004.
- [13] LiClipse, Brainwy Software Ltda., [En línea]. Disponible: <http://www.liclipse.com/>. [Último acceso: 4 febrero 2016].
- [14] CMU-SPHINX, "sourceforge.net", Carnegie Mellon University, [En línea]. Disponible: <http://cmusphinx.sourceforge.net/wiki/tutorialam>. [Último acceso: 3 febrero 2015].
- [15] SoX Sound eXchange, [En línea]. Disponible: <http://sox.sourceforge.net/>. [Último acceso: 4 febrero 2016].

# Diseño e instrumentación de un sistema de monitoreo de pulso y temperatura en bebés vía Wi-Fi

Miguel Angel Ortega  
Benemérita Universidad Autónoma de Puebla  
Faculty of Electronic, Mexico  
Email: miguelortpal@ece.buap.mx

Sergio Vergara  
Benemérita Universidad Autónoma de Puebla  
Faculty of Electronic, Mexico  
Email: svergara2@ece.buap.mx

Aurora Vargas  
Benemérita Universidad Autónoma de Puebla  
Faculty of Electronic, Mexico  
Email: auroravargast@hotmail.com

David Pinto  
Benemérita Universidad Autónoma de Puebla  
Faculty of Computer Science, Mexico  
Email: dpinto@cs.buap.mx

Darnes Vilariño  
Benemérita Universidad Autónoma de Puebla  
Faculty of Computer Science, Mexico  
Email: darnes@cs.buap.mx

**Resumen:** El síndrome de muerte súbita del lactante, también conocido como muerte de cuna, es la muerte repentina e inesperada de un bebé menor de un año, sin razón aparente. Es más común en los primeros meses de vida. Se conoce poco acerca de la causa, pero se relaciona con problemas en el mecanismo del sistema nervioso central que controla la respiración y el ritmo cardíaco; tal mecanismo está involucrado en la detección de intercambio de gases durante la respiración; además se presenta de forma esporádica y rápida. Regularmente no se puede prevenir o evitar. Para contrarrestar esta situación, es necesario diseñar y construir un dispositivo de tobillo que permita detectar los parámetros biomédicos (temperatura corporal y ritmo cardíaco) del bebé desde cualquier dispositivo móvil o computadora, a partir de información transmitida inalámbricamente mediante la interfaz Wi-Fi, la cual estará recibiendo la información en tiempo real de la temperatura y el pulso proporcionados por los sensores optoelectrónicos. Los padres o el pediatra serán notificados mediante un aviso en el momento que se presente alguna alteración en los signos vitales del bebé.

**Abstract:** Sudden Infant Death Syndrome, also known as crib death, is the sudden and unexpected death of an infant under one year for no apparent reason. It is more common within the first months of age. Little is known about the cause, but is related to problems in the central nervous mechanism system that controls breathing and heart rate; such mechanism is involved in the detection of gas exchange during respiration; besides it occurs sporadically and quickly. Regularly it cannot be prevented or avoided. To counter this situation, it is necessary to design and build an ankle device for monitoring biomedical parameters (body temperature and pulse rate) of infant from any mobile device or computer, from data transmitted wirelessly by the Wi-Fi interface, which will be receiving real-time temperature and pulse provided by the optoelectronic sensors. Parents or pediatrician will be notified with a warning at the time that an alteration in the baby's vital signs is present.

*Keywords:* Sudden Infant Death Syndrome; body temperature; pulse rate; Wi-Fi.

## I. INTRODUCCIÓN

El síndrome de muerte súbita del lactante se define como el fallecimiento repentino e inesperado de un niño menor de un año aparentemente sano, que permanece sin explicación después de una investigación exhaustiva del caso, incluyendo la realización de una autopsia completa, el examen de la escena de la muerte, y la revisión del historial clínico. Ocupa uno de los primeros lugares como causa de muerte durante el primer año de vida, con un pico entre los dos y tres meses de edad. En países desarrollados se ha estimado que hasta 55 % de las muertes en niños menores de un año se debe a este síndrome. En Estados Unidos de Norteamérica ocupa el tercer lugar como causa de muerte, después de las malformaciones congénitas y el peso bajo al nacer asociado con la gestación corta, con una tasa de 0.57 muertes por mil nacidos vivos en 2002, después de una reducción de 53 % en 10 años. En Latinoamérica, la tasa varía entre 0.1 y 5.4; en el año 2000 en México, la tasa se estimó en 0.6. Hasta el momento no se conoce la causa ni el mecanismo de cómo se produce la muerte en este síndrome, sin embargo, se han identificado factores de riesgo relacionados con la madre, el niño y las condiciones en las que duerme el niño [1-2].

## II. MEDICIÓN DEL PULSO EN UN BEBÉ

El pulso es el número de latidos del corazón por minuto. Con cada latido el corazón bombea sangre a todo nuestro cuerpo con las sustancias necesarias para mantener nuestras células saludables. Si el corazón no funciona correctamente, la persona corre serios riesgos de salud. La medición del pulso proporciona información importante acerca del estado de salud de una persona. Existen varias maneras de medir las pulsaciones de una persona. En niños menores de un año, se recomienda tomar el pulso radial o pedio. En niños mayores al año, se pueden utilizar las arterias que pasan por la muñeca, el cuello (carótida) o la



ingle. Es importante mencionar que no se debe tomar el pulso con el pulgar, los dedos que se utilizan son el anular y el mayor juntos [3].

La frecuencia normal del pulso varía con la edad y con su estado clínico, así a menor edad mayor frecuencia del pulso. El incremento en la frecuencia del pulso puede indicar excitación, fiebre, insuficiencia cardíaca congestiva o disritmias. La bradicardia puede ser un signo de bloqueo cardíaco. Un pulso irregular puede ser sugestivo de disritmias pero se debe tener cuidado de no confundirla con arritmia sinusal en la que se produce una aceleración de la frecuencia cardíaca con la inspiración, lo cual es normal y acentuado en los niños. En los neonatos prematuros los pulsos son saltones por la disminución del tejido subcutáneo y porque muchos tienen asociada una persistencia del conducto arterioso [4].

### III. HARDWARE

En la Fig. 1 se muestra el diagrama general del sistema de monitoreo, en donde se representa el funcionamiento principal del mismo y los segmentos más importantes que lo conforman.

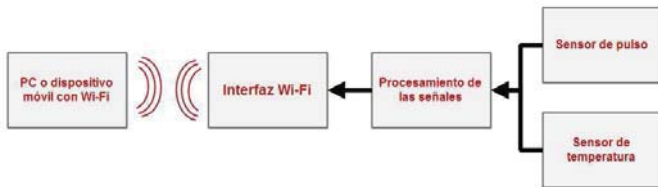


Fig. 1. Diagrama a bloques del hardware del sistema.

Se instrumentó un sensor optoelectrónico de pulso para medir el ritmo cardíaco a través de un arreglo de leds infrarrojos y fotodiodos. La cantidad de fotocorriente generada por el sensor se basa en la luz incidente y la longitud de onda. La primera parte del circuito consta de un amplificador de transimpedancia, cuya función es hacer una conversión de la cantidad de luz incidente a voltaje, en función de la corriente que proporciona el fotodiodo. Para eliminar las componentes constantes de la señal a sensar, se propuso la implementación de un filtro pasa altas con una frecuencia de corte ( $f_c$ ) de 0.5 Hz. El filtro pasa altas tiene como finalidad filtrar la señal de componente continua y solo permite adquirir señales de frecuencias superiores a la frecuencia de corte. Debido a que el ritmo cardíaco de un bebé puede oscilar de entre 120 a 180 pulsos por minuto aproximadamente [5-6], se propuso la implementación de un filtro pasa bajas para evitar señales no deseadas, el circuito tiene como finalidad limitar la frecuencia de trabajo, ya que la señal se caracteriza por presentar componentes relevantes hasta la frecuencia de corte, por lo que se propuso una frecuencia de corte aproximada de 4 Hz. Finalmente, la última parte del circuito consta de una etapa de amplificación, empleando un amplificador inversor con una ganancia variable. El funcionamiento del circuito completo del sensor optoelectrónico de pulso se muestra en el diagrama a bloques de la Fig. 2.



Fig. 2. Diagrama a bloques del circuito para el sensor optoelectrónico de pulso.

Para la medición de la temperatura corporal, se propuso utilizar el sensor MLX90614, el cual es un termómetro infrarrojo para mediciones de temperatura sin contacto. Tanto la termopila sensible al infrarrojo y el acondicionamiento de la señal están integrados en el mismo encapsulado. Gracias a su amplificador de bajo ruido, convertidor analógico-digital de 17 bits y potente unidad DSP (Digital Signal Processing), el termómetro tiene una alta precisión y resolución. El termómetro viene calibrado con salida PWM (Pulse Width Modulation) digital y SMBus (System Management Bus) de fábrica. Como estándar, el PWM de 10-bits está configurado para transmitir de forma continua la temperatura medida en el rango de  $-20$  a  $120$  °C, con una resolución de salida de  $0.14$  °C. El ajuste predeterminado de fábrica es SMBus [7].

Se ejecutaron pruebas de medición de la temperatura corporal para observar el comportamiento de la respuesta del sensor con respecto a la distancia que se encontraba de la piel, la medición se hizo para medir la temperatura corporal en una mano. La curva característica de la temperatura en función de la distancia del objeto se aprecia en la Fig. 3. Como se puede observar, la gráfica tiene un comportamiento exponencial, y a menor distancia la medición del sensor es más precisa. Además se pudo comprobar que a una distancia de aproximadamente 1 cm, la lectura de la temperatura es la ideal.

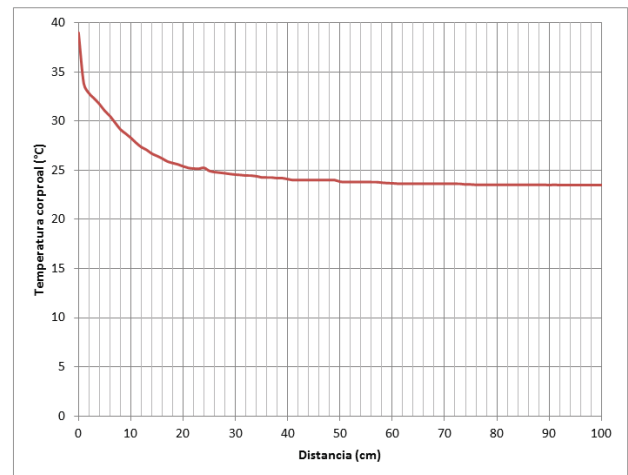


Fig. 3. Gráfica de la medición de la temperatura corporal con respecto a la distancia que se encuentra el sensor de la piel.

Una vez comprobado el funcionamiento del dispositivo, se prosiguió a realizar la placa de circuito impreso (PCB) donde se colocaron ambos sensores, el de pulso y temperatura, de tal forma que se pudiera tener ambos sensores en la misma placa. Adicionalmente se propuso utilizar una placa doble cara para reducir su tamaño y distribuir de mejor forma cada uno de los dispositivos electrónicos. Se utilizó el software Altium Designer para hacer la PCB, el cual se aprecia en la Fig. 4.

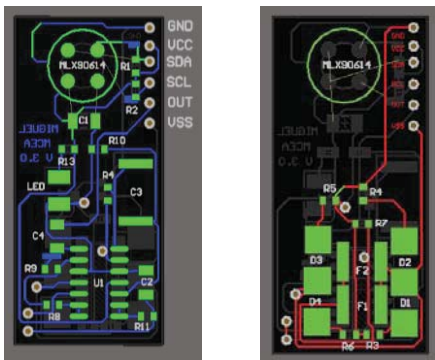


Fig. 4. Capa inferior y superior de la placa de circuito impreso.

A partir del diseño de la PCB, se diseñó la placa, y el resultado final se identifica en la Fig. 5. Como se visualiza, los leds infrarrojos se colocaron muy cerca de los fotodiodos, ya que de esta manera, se mejora el desempeño del dispositivo y es más fácil detectar la señal del ritmo cardiaco. Además, la placa fue cubierta con foami para cubrirla de ambos lados, dejando a la vista el sensor de temperatura, y el arreglo de fotodiodos y leds infrarrojos, para que puedan estar en contacto directo con la piel.



Fig. 5. Diseño final de la placa para el sensor de pulso y temperatura.

#### IV. FIRMWARE

Por otro lado, se generó el respectivo código de programación para la lectura de los sensores, el cálculo del umbral variable y la frecuencia cardiaca; para ello se empleó una tarjeta Arduino Pro Mini, debido a que es muy pequeña. El funcionamiento básico del firmware se simplifica en el diagrama a bloques de la Fig. 6.

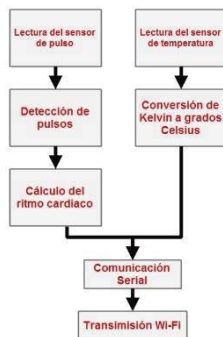


Fig. 6. Diagrama a bloques del firmware del sistema.

Para detectar el momento en el que se presenta un pulso se tomaron en cuenta tres muestras de la señal del sensor, que corresponden a la muestra anterior, la muestra actual y la muestra siguiente. De tal manera que si la muestra actual es mayor a la muestra anterior y a la muestra siguiente, entonces se detectará un pico en la señal. Posteriormente, cada vez que un pico sobrepase el valor de umbral, se contará un pico y se guardará la cuenta en una variable.

Se propuso un nivel de umbral variable puesto que la amplitud de los pulsos no es siempre la misma, entonces se empleó un filtro de respuesta de impulso infinito, este tipo de filtro se aplica para eliminar ruido de alta frecuencia [8]. Se propuso que el valor de umbral actual estuviera en proporción al 80% del valor de umbral anterior y la mitad del 20% del valor máximo de la muestra anterior. Se le asignó mayor valor al umbral anterior para que el valor del umbral actual no tuviera cambios muy abruptos al momento de actualizar su amplitud, la ecuación para generar el umbral se obtiene de (2), la cual se genera a partir de (1).

$$y[i] = \frac{1}{k} \{x[i] + (k - 1)y[i - 1]\} \quad (1)$$

$$Umbral = 0.8(Umbral) + 0.1(Muestra actual) \quad (2)$$

Se realizaron mediciones del ritmo cardiaco en una persona adulta durante un lapso de 60 segundos, para comprobar el comportamiento del umbral variable a partir de los pulsos detectados y se graficaron los resultados, los cuales se muestran en la Fig. 7, en color verde se observa la señal obtenida del sensor de pulso, en color rojo se muestran los pulsos detectados a partir de la condición del nivel de umbral; en total se detectaron 75 pulsos en un minuto. Y en color azul se identifica el nivel de umbral, el cual cambia su valor con respecto a la amplitud de la señal del sensor.

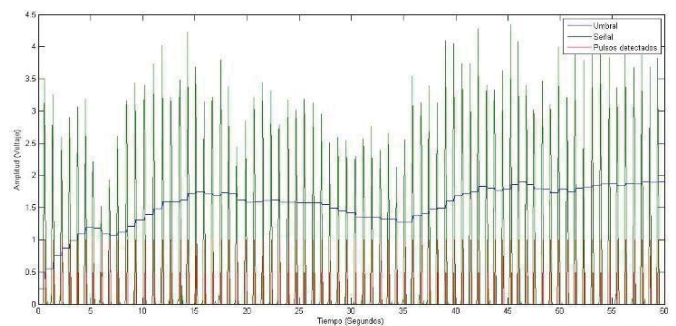


Fig. 7. Comportamiento del umbral variable con respecto al tiempo.

En la Fig. 8 se hace un acercamiento a la gráfica de la Fig. 7, para visualizar mejor el comportamiento de la detección de pulsos y el nivel de umbral variable, durante un lapso de 10 segundos aproximadamente.

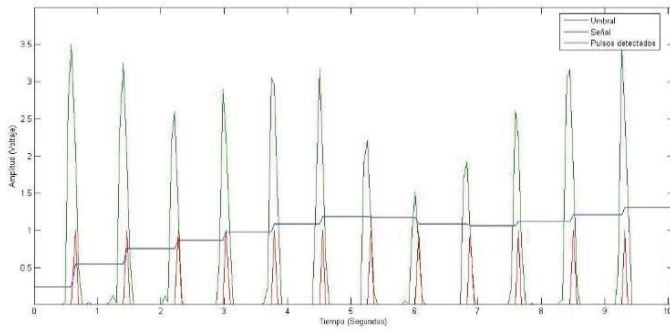


Fig. 8. Acercamiento a la gráfica de la respuesta del umbral variable con respecto al tiempo.

Para obtener el promedio de pulsos por minuto, se contabiliza el número de pulsos y el tiempo en el que estos ocurren y posteriormente se aplica un filtro de respuesta de impulso infinito para mejorar su desempeño, como se muestra en (3). Dándole mayor porcentaje al valor del promedio actual que al anterior para evitar cambios abruptos.

$$Pulso = 0.9(Pulso\ actual) + 0.1(Pulso\ anterior) \quad (3)$$

En la Fig. 9 se hace únicamente la comparación entre la respuesta del promedio del ritmo cardiaco obtenido y la señal del sensor de pulso, durante el lapso de un minuto. Tal como se valida, el comportamiento del promedio presenta un pequeño sobreimpulso y tarda aproximadamente 15 segundos en estabilizarse, posteriormente se mantiene cerca del valor de 75 pulsos por minuto, por lo tanto se comprueba el buen desempeño del algoritmo propuesto.

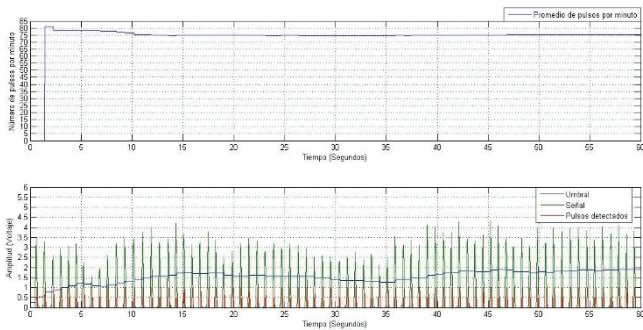


Fig. 9. Comparación de la respuesta del promedio del ritmo cardiaco y la señal del sensor de pulso, con respecto al tiempo.

## V. COMUNICACIÓN WI-FI

Para realizar la comunicación inalámbrica del sistema con cualquier dispositivo móvil o PC, se propuso utilizar el módulo ESP8266, que ofrece una solución completa y muy económica para conexión de sistemas a redes Wi-Fi, es pequeño (24.8 x 14.3 mm), y tiene un buen nivel de prestaciones para ser usado con microcontroladores de gama media y alta, el módulo se puede identificar en la Fig. 10.

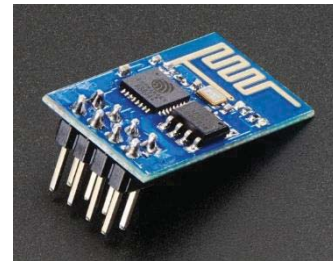


Fig. 10. Módulo ESP8266.

El módulo ESP8266 es capaz de funcionar como adaptador de red en sistemas basados en microcontroladores que se comunican con él a través de una interfaz UART [9].

Para adaptar el módulo al sistema de monitoreo, se diseñó una pequeña placa para montarlo y se agregó un regulador de voltaje ajustable a 3.3V, utilizando el circuito integrado LM1117, debido a que si el módulo ESP8266 se alimenta a mayor voltaje, se puede sobrecalentar y dañar.

El siguiente paso fue realizar algunas pruebas en la página web llamada ThingSpeak para enviar los datos que proporcionan los sensores de pulso y temperatura. ThingSpeak es una plataforma abierta de aplicaciones, diseñada para permitir conexiones significativas entre las personas y los objetos. Se trata de una aplicación Open Source para la Internet de las Cosas y una API (Application Programming Interface) para almacenar y recuperar datos de los objetos usando HTTP sobre Internet o a través de una red de área local. Con ThingSpeak, se pueden crear aplicaciones de sensores de registro, aplicaciones de seguimiento de localización, y una red social de las cosas con las actualizaciones de estado. Puede funcionar a través de un servicio gratuito de host, o bien, en servidores personales. La información es procesada y presentada de un modo significativo. Toda la información es abierta y está accesible para que otros desarrolladores puedan hacer uso de ella [10].

La única limitante que se tiene es que la página se actualiza cada 15 segundos, por lo tanto cada 15 segundos se puede visualizar un dato nuevo en la gráfica.

Primeramente, es necesario crear los campos que se van a requerir para el monitoreo, uno para graficar la temperatura corporal (°C) y otro para graficar la frecuencia cardiaca (Número de pulsos por minuto), así como se aprecia en la Fig. 11.



Fig. 11. Diseño de los campos para graficar la temperatura corporal y la frecuencia cardiaca.

## VI. RESULTADOS EXPERIMENTALES

Para la alimentación del sistema se utilizaron tres baterías de tipo Polímero de Litio (Li-Po), las cuales proporcionan un voltaje 3.7V cada una, con una celda por batería y capacidad diferente, una de 850mAh, otra de 550mAh y la última de 240mAh. Las baterías de 850mAh y 550mAh se conectaron en serie para poder tener un voltaje resultante de 7.4V y poder obtener el voltaje regulado de 3.3V a partir de la salida del LM1117. Y posteriormente, el arreglo de estas baterías se conectó en serie con la batería de 240mAh para obtener el voltaje negativo de -3.7V para alimentar el sensor de pulso. Las 3 baterías que se utilizaron para la alimentación del sistema completo, se identifican en la Fig. 12.

Finalmente se fijó cada una de las partes que conforman el sistema completo para realizar un prototipo final de la pulsera, el cual se colocó en un tramo de venda flexible de aproximadamente 20 cm. de largo y 5 cm. de ancho, el cual se puede ajustar con velcro, dependiendo del tamaño que se requiera. En la parte interna de la pulsera se colocaron los sensores para que pudieran estar en contacto directo con la piel, y en el exterior de la pulsera se colocó el resto de los dispositivos (ESP8266, baterías, Arduino Pro Mini); el prototipo final se muestra en la Fig. 12 y Fig. 13.

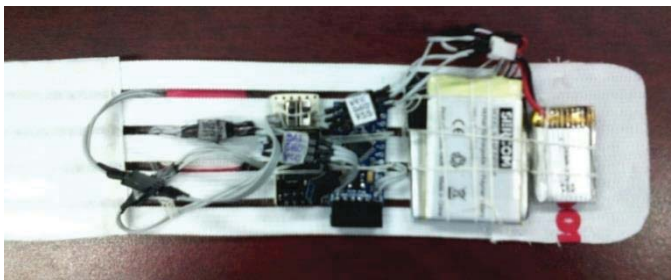


Fig. 12. Parte externa de la pulsera.



Fig. 13. Parte interna de la pulsera.

Una vez terminado el diseño de la pulsera, se hicieron pruebas finales para comprobar el funcionamiento del sistema completo. Se colocó la pulsera en la muñeca izquierda de una persona adulta para hacer mediciones de su temperatura corporal y frecuencia del pulso radial, tal como se muestra en la Fig. 14.



Fig. 14. Colocación de la pulsera para medir la temperatura corporal y la frecuencia del pulso radial en una persona adulta.

Se realizaron varias pruebas en la página ThingSpeak para observar el comportamiento de la señal de ambos sensores, enviando un dato cada 15 segundos, graficando la temperatura corporal (°C) y el ritmo cardiaco (Pulsos por minuto) durante un periodo de tiempo. En total se ejecutaron cuatro pruebas diferentes para medir la temperatura corporal del sujeto y su frecuencia cardiaca, a partir de su pulso radial. Comparando los resultados de la Fig. 15, Fig. 16, Fig. 17, Fig. 18, se puede validar que la temperatura se mantiene constante y en un valor cercano a los 36°C para los cuatro casos. Sin embargo, la frecuencia cardiaca presenta pequeñas variaciones a través de tiempo, en algunos lapsos se encuentra aproximadamente en los 75 pulsos por minuto y en otros lapsos se mantiene arriba de este valor. Algunos picos que se aprecian en las gráficas son debido a que se presentó movimiento en la extremidad donde estaba colocada la pulsera, por lo que se alteraba la señal.



Fig. 15. Prueba 1 en ThingSpeak, utilizando la pulsera para medir la temperatura corporal y la frecuencia del pulso radial.

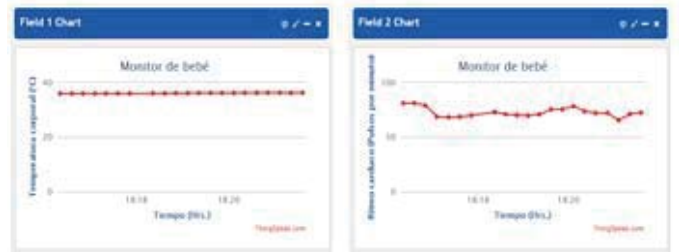


Fig. 16. Prueba 2 en ThingSpeak, utilizando la pulsera para medir la temperatura corporal y la frecuencia del pulso radial.

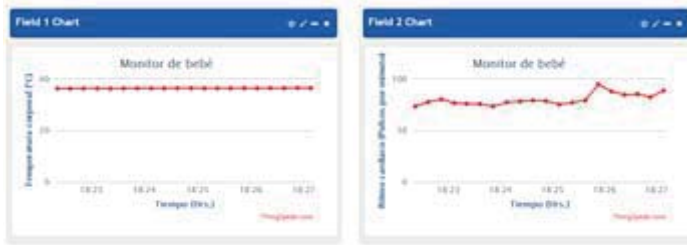


Fig. 17. Prueba 3 en ThingSpeak, utilizando la pulsera para medir la temperatura corporal y la frecuencia del pulso radial.



Fig. 18. Prueba 4 en ThingSpeak, utilizando la pulsera para medir la temperatura corporal y la frecuencia del pulso radial.

Por lo tanto, es importante mencionar que durante las mediciones, es necesario mantener sin mucho movimiento la extremidad en donde se coloca la pulsera, de lo contrario, se presentará una respuesta no deseada en la medición de la frecuencia del ritmo cardiaco.



Fig. 19. Implementación del prototipo final para medir la temperatura corporal y la frecuencia del pulso radial.

## VII. TRABAJO A FUTURO

Para concluir el proyecto, es necesario reducir el tamaño del sistema completo, debido a que deberá ser colocado en el tobillo de un bebé menor a un año. Por lo que probablemente se hará un cambio en el diseño de la placa de los sensores para reducir su tamaño y es importante reemplazar las baterías por unas de menor tamaño y mayor capacidad. Esto será de gran ayuda para reducir el tamaño del dispositivo. Y finalmente se implementará un sistema de alarma para el sistema de monitoreo para avisar a los padres o al pediatra, a través de algún dispositivo móvil o

computadora, en el momento que se presente una alteración en los signos vitales del bebé.

## REFERENCIAS

- [1] S. Flores, R. I. Ramos, S. Flores, S. Villa, H. Martínez, “*Síndrome de muerte súbita del lactante. Prevención en la práctica hospitalaria*”, Rev Med Inst Mex Seguro Soc 2006, p. 511.
- [2] G. Ottaviani, “*Crib Death. Sudden Unexplained Death of Infants: The Pathologist’s Viewpoint*”, p. 2, 2007.
- [3] M. I. Soldano, “*Primeros Auxilios Y Enfermedades de la Infancia*”, Primera Ed. p. 13, Octubre 2005.
- [4] A. Marín, J. F. Gómez, J. C. Jaramillo, “*Manual de pediatría ambulatoria*”, 21 ed., pp. 356-358, 2008.
- [5] P. J. Schwartz, A. Garson, Jr, T. Paul, M. Stramba-Badiale, V. L. Vetter, E. Villain and C. Wren, “*Guidelines for the interpretation of the neonatal electrocardiogram*”, Vol. 23, p. 1331, September 2002.
- [6] B. J. Deal, C. L. Johnsrude, S. H. Buck, “*Pediatric ECG Interpretation: An Illustrative Guide*”, p. 8, 2004.
- [7] Data Sheet MLX90614 family, “*Single and Dual Zone Infra Red Thermometer in TO-39*”, pp. 1-2, February 28, 2013.
- [8] AN85951 – *PsoC 4 CapSense Design Guide*, Cypress Semiconductor Corporation, Doc. No. 001-85951 Rev. M, p. 49, 2013-2016.
- [9] Espressif System, “*ESP8266 802.11 bgn Smart Device*”, pp.5-6, October 12, 2013.
- [10] Hans Scharler, “*ThingSpeak: Open Source Platform for Connected Products and Services*”, ioBridge, Inc., April 18, 2011.